United Nations ECE/CES/2014/24

Economic and Social Council

Distr.: General
28 January 2014
English
Original: French

Economic Commission for Europe

Conference of European Statistians

Sixty-second plenary session
Paris, 9–11 April 2014
Item 3 of the provisional agenda
What is the value of official statistics and how do we communicate that value?

How do we communicate the value of official statistics?

Prepared by the Institut national de la statistique et des études économiques (National Institute of Statistics and Economic Studies), France

Summary

To communicate the value of official statistics we must communicate information on the factors that go into them, specifically their relevance and the confidence that people have in the data. The communication must target priority groups, using appropriate means. In France, the priority groups are the media, government bodies and the public. The communication strategy of the Institut national de la statistique et des études économiques (National Institute of Statistics and Economic Studies) (INSEE) is built on three pillars: education, connection and proof. It communicates the value of the statistics mainly by means of the following: an annual report; a communication toolkit, issued yearly; answers to queries; and meetings, several of which have been scheduled for late 2013 and early 2014 on the theme, "official statistics and democracy: what good are figures?"

This document is being submitted for discussion at a seminar of the European Conference of Statisticians on the question, "What is the value of official statistics and how do we communicate that value?"

GE.14-20438 (E) 240214 260214

1420438

Please recycle

I. Introduction

1. With the Internet, an increasing amount of data is being distributed. Since Web 2.0, anyone can produce data and send it to everyone through the web. As "big data" has developed, companies and bodies have been able to analyse all these various types of information and produce synthetic data. There is a risk that official statistics may be lost in this immense mass of data, as people will not know about them or will have faith in unofficial statistics.

2. Official statistics have a value that sets them apart from other data. They are constructed according to methods and guidelines that give them an upper hand (in terms of reliability and comparability, for instance) and that make it possible to use them as reference data.

4. Statisticians have to be able to assess the value of official statistics and to communicate that value. This document gives some indications on how this can be done.

5. Before discussing how to communicate the value of official statistics, we should go over what the value is. According to recent studies, the value (or added value) of official statistics (V) can be measured as follows:

V = S * M * A * R * T * L, where:

S is the number of statistics produced;

M is a measure of the performance of the media used to reach the target audience;

A is the number of persons who require the official statistics, or the potential audience. The product of M and A gives the number of users who find the official statistics;

R is the relevance of the official statistics sent to users. From the users' standpoint, this is the quality of the statistics;

T is the trust in the official statistics; and

L is statistic and economic literacy[1] and culture, or the ability to reason using numbers and other mathematical concepts, and using economic concepts.[2] It reflects the ability of users to properly understand, interpret and use statistics.

6. The value of official statistics thus depends on these six factors. If one factor is absent, the value is nil.

II. What can we communicate about the value of official statistics?

7. In this part we will take a look at what we can use to communicate the value of official statistics. If we want to get across the value of statistics, we cannot directly speak of

[1] According to the Organisation for Economic Co-operation and Development (OECD), "literacy" is defined as "the ability to understand and employ printed information in daily activities, at home, at work and in the community – to achieve one's goals, and to develop one's knowledge and potential".

[2] Statistical culture is the ability to comprehend numbers, quantities, basic mathematical concepts, the main statistical indicators (medians, averages, proportions), statistical reasoning and the meaning of a probability.

"the value of statistics" per se. We have to address one or more of these six factors: statistics, media, potential audience, relevance, trust and literacy (SMARTL). We can also communicate the value by referring to elements that consolidate these six factors. For example, we can mention the independence of the body issuing the statistics, which consolidates trust in it.

8. (S) We can provide information on the number of official statistics produced and distributed. This is not so easy to do, as there are very many official statistics. To do so, we can for example provide information on:

(a) Subjects covered (unemployment, population, poverty, etc.);

(b) Kinds of data distributed (tables with main indicators, databases, etc.); and

(c) Volume of data distributed (in megabytes or gigabytes).

9. (M) We can provide information on the media that we use to distribute our data, as follows:

(a) General media: our websites, publications, smartphone apps, web applications and social media (Twitter, etc.);

(b) Our specialized media: for journalists, our press conferences and press releases; for expert users, conferences and meetings, etc.; and

(c) Media of other organizations: other websites (data.gouv, Ministry of Education, partners' websites), press, exhibitions, museums, etc.

10. (A) It is difficult to measure and give information on the potential audience. However, we can provide information on the actual audience, which is the product of the media performance, M, and the potential audience, A (or M * A). For example, it is possible to communicate:

(a) Some figures about visits to our website: number of visits, of visitors, or of gigabytes downloaded (daily, monthly or annually); and

(b) Some figures about social media: number of subscribers to our Twitter account, etc.

11. (R) To provide information on the relevance of official statistics, first we must measure it, which means collecting our users' opinions. With the Internet, this is not easy, as more often than not we do not see our users. In order to be effective, and because collecting users' opinions helps first to improve the statistics and later to improve communication, we should give preference to:

(a) Collecting their opinions on specific subjects, and on new subjects every year;

(b) Working with colleagues who are specialists in the selected subjects;

(c) Interviewing different users: specialists and beginners, and if possible, also potential users, teachers, local authorities, people in national government, researchers, students, journalists, private companies, associations and individuals; and

(d) Using one or several methods to collect their opinions in relation to the subject in question; online survey using a pop-up window on our website; online e-mail survey; focus groups; personal interviews; phone interviews; analysis of the queries we receive; analysis of consultation statistics for our website, etc.

12. (T) To provide information on the trust that people have in official statistics, we should first measure it by asking people about their trust in statistics. It is important to interview at least two groups every year: the public at large, and visitors to the INSEE

website, as the two groups give different opinions; website visitors are more specialized than the public. It is also possible to interview other groups when necessary. For example, a group of journalists may be interviewed if a specific issue concerns them or if the Institute's strategy involves specific activities for them. In order for such interviews to help improve the statistics, and to be able to communicate several of the constituent elements of trust, at least the following questions should be asked:

(a) A general question about statistics, such as, "Generally, would you say that you trust the figures and data published on the economic and social situation in France?";

(b) If people find the figures and data published on the economic and social situation in France to be suspect, then some questions should be asked to find out why, for example to elicit replies such as: "These figures have no correspondence with my situation and the situation of those around me"; "Politicians make them say what they want them to say"; or "I do not know what these figures represent or how they are produced";

(c) Subsequently, a series of individual indicators should be addressed, including the unemployment rate, the growth rate and the consumer price index, with the question: "For each of the following figures and data, please tell me if they give a very good, good, poor or very bad reflection of what is happening in France in the field in question"; and

(d) At the end, questions should be posed regarding the Institute: a general question such as, "Is your opinion of INSEE very good, good, bad or very bad", followed by questions such as, "Is INSEE a useful body?", "Does INSEE provide good information on its activities?", "Is the information produced and disseminated by INSEE independent of political influence?" and "Is the information on the French economy produced and distributed by INSEE reliable?"

13. (L) It is very difficult to measure statistical and economic literacy. We are currently carrying out a survey regarding some definitions (median, gross domestic product (GDP), etc.) and will have the results in a few months. There may then be something to communicate on this subject.

14. Among the six factors, communication is easy for the first three (the number of statistics, the media used and the audience) and very difficult (or even impossible) for the last (literacy). It is not easy, but very important, to provide information on the fourth and fifth factors: relevance and trust.

15. V=SMARTL is a general formula, and it can also be used for a specific audience. The value of official statistics is not the same for the general public and for expert audiences, as their needs are not the same. The relevance factor, R, is dependent on the users' needs.

16. The formula can also be used for one part of official statistics, for example the statistics covering a given subject or a specific indicator, such as housing statistics or the unemployment rate.

III. How to communicate the value of official statistics?

17. In the preceding part we looked at the possible content to be used in communicating the value of official statistics. In this part we will look at how to communicate this content, giving a few examples of how this has been done in France.

18. How to communicate the value of official statistics depends on the overall communication strategy of the statistics body, which decides which audiences and indicators should be given priority.

19. In France, INSEE selects the priority audience based on its importance and its opinions of statistics. INSEE has three priority audiences: the media, local and national government figures and the public at large. It is indeed necessary to maintain good relations with the media and the authorities. Because opinions of statistics are much more negative among the general public than in other groups, five years ago INSEE added the public as a priority audience. INSEE decides which indicators to prioritize by combining two criteria: the most general and generally accessible, and the ones in which there is the lowest level of confidence. The consumer price index and the unemployment rate are among the indicators that are thus given the highest priority.

20. Once the target audiences and the subjects have been selected, another decision remains: what media to use. The media depends on the target audience, the subjects and the overall communication strategy of the Institute. In France, in order to give the indicators a stronger image, INSEE has drawn up a list of action to be taken in the short and medium term, in line with these three factors:

(a) Connection: People who are connected with the Institute have a better image of its work and of the indicators, and many people wish to know more about the Institute's activities;

(b) Education: One of the elements that explain mistrust of indicators is that people do not know how they are produced; and

(c) Proof: Another factor breeding mistrust of indicators is that people think they do not correspond with reality. Many people would like these figures and data to better reflect their personal situation and the situation of the people in their entourage.

21. The value of statistics is communicated, for example, by means of: the Institute's annual report; a communication toolkit; discussions at meetings or conferences; teaching tools; and through the use of the website.

22. The annual report provides a good means of highlighting the value of statistics, especially for the local and national authorities, by addressing five of the six factors in SMARTL: the kind of statistics disseminated; the main and new kinds of media used in the process; the audience of our website and Twitter account; and the most important results of our study on relevance and trust. For example, in our 2012 annual report we wrote the following:

(a) On www.insee.fr, users can now find millions of data of many kinds: framework data and key figures; historical data and time series; localized data for regions, municipalities, and neighbourhoods; and microdata at individual level, anonymized to comply with statistical confidentiality requirements [...];

(b) www.insee.fr has thus become INSEE's dissemination medium of choice. The number of visits has risen steadily from 10.8 million in 2004 to 25 million in 2012. However, while surveys indicate high satisfaction with the wealth of information available, visitors still report trouble locating the data they need. We are therefore preparing a new redesign, notably to improve navigation [...];

(c) 120,000 visits to www.insee.fr daily, 30,000 subscribers to *INSEE-Net Actualité*;

(d) 10,000 Twitter subscribers and 11 press conferences in 2012 [...];

(e) The press and audiovisual media are a vital link for the dissemination of INSEE statistics and studies. On average, our work is referred to in 60 media stories a day.

23. Every year, INSEE issues a kit on a different indicator (consumer price index, unemployment rate, growth rate, etc.) to show the public how it constructs the indicator,

who uses it and for what purpose. The kit is composed of three communication tools: an educational brochure entitled *Pour comprendre …* (To understand …); a short (under five-minute) video entitled *… en questions* (… in questions); and a quiz composed of 10 multiple-choice questions whose answers are in the brochure and video. Among other things, these tools make it possible to communicate the relevance of the data. We also use Twitter to tweet messages on the indicators and main results, and also on the tools used.

24. In 2013 the Conseil national pour l'information statistique (National Council for Statistical Information) (CNIS) and INSEE held several activities to strengthen confidence in official statistics. CNIS provides a forum for consultations between producers and users of official statistics, its activities ranging from the identification of sectors and subjects to be covered by surveys, to evaluations of surveys once they are completed. It has its own publications to inform the public. The Council has therefore held several activities with INSEE aimed at strengthening confidence in official statistics, under the general theme of "Official statistics and democracy: what good are figures?" The main purpose of these activities is to make the media, the local and national authorities and the general public aware of the value of official statistics.

25. Four activities have been carried out:

(a) Every year an Economics Day is held in November in Lyons, with the participation of a broad range of participants. In November 2013 INSEE and CNIS experts used studies of well-being to demonstrate the value of official statistics;

(b) In December 2013, a "media and statistics" seminar was held with the School of Journalism in Lille. CNIS and INSEE took part in this seminar in order to show future journalists the value of official statistics at a time when data journalism is developing;

(c) INSEE worked with an association called les Petits Débrouillards, which specializes in promoting science with children and teenagers, to develop teaching tools that will be used throughout France and will help strengthen confidence in official statistics; and

(d) The culmination of this approach will be a national conference and discussion scheduled for the end of January 2014. A broad range of statistics users have been invited: researchers, decision makers, public figures and journalists. The theme of the conference is to be "the usefulness of statistics".

26. We intend to put more information on the value of our statistics on our website in order to strengthen confidence in our statistics and their relevance. We intend for example to:

(a) Develop a section of our website to give the main results of satisfaction surveys conducted with users of official statistics;

(b) On our home page, communicate information that is not statistical, but that relates to the value of statistics, specifically the elements mentioned in paragraphs 17 to 20, above; and

(c) Include in some publications or press releases accounts from decision makers demonstrating in one or two simple sentences how they use the results and indicators that we publish.

27. Paragraphs 17 to 21 give examples of communication activities that we have decided to carry out and have prepared.

28. Furthermore, some communication activities provide immediate (or almost immediate) responses to a question or event. They can thus prevent a discussion that might challenge the quality of statistics in general or of our Institute or a given indicator from turning bad and from undermining confidence in official statistics.

29. We rely mainly on three means of following discussions relating to statistics and our Institute: by monitoring the Internet and social media; the press; and incoming questions. When selecting how to monitor the Internet and social media, every year we choose websites of interest, which cover our subjects and have audiences that are qualitatively and quantitatively significant. We do the same when we select which media outlets to monitor.

30. In general, we do not respond to questions or events found through our monitoring, as it is often more risky to respond than not to respond. When it is necessary to respond to a question, we do so through the appropriate media: press release, tweet, comment on one of our users' blogs, personal mail or message or phone call, etc. It is not easy to know when to reply, nor which media is most appropriate. We intend to analyse a number of cases handled over several months so as to draw up a strategy.

V. References

Giovaninni, Enrico (2008). *The role of communication in transforming statistics into knowledge*. http://unstats.un.org/unsd/accsub/2008docs-CDQIO/Ses3-Pap2.pdf.

Novkovska, Blagica (2011). *Building a basis for measurement of the value added in official statistics*. http://www.ksh.hu/docs/eng/xftp/idoszaki/visegrad.pdf.

Mitton, Terri (2013). *OECD's approach to understanding user needs*. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.45/2013/Session_3_-_Terri_Mitton.pdf.

INSEE (2013). *2012 Annual Report*. http://www.insee.fr/en/insee-statistique-publique/connaitre/rae/en-rae.pdf.

Brunet, François (2013). *Ten recommendations to collect feedback from users*. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.45/2013/Session_3_-_Francois_Brunet.pdf.

United Nations ECE/CES/2014/25

Economic and Social Council

Distr.: General
30 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 3 of the provisional agenda
What is the value of official statistics and how do we communicate that value?

Communicating innovation in National Statistical Offices

Prepared by the National Institute for Statistics of Italy*

Summary

Modernisation strategies for National Statistical Institutes (NSIs) are driven by four key factors. These include standardisation and industrialisation of the statistical business chain, implementation of re-usable infrastructure, adoption of metadata-driven systems and development of client-oriented communication strategies.

As part of NSI innovation strategies, statistical data need to be proactively communicated to engage users through a customer-centric approach. Innovative communication tools can help respond to increasing and changing user needs.

NSIs need to be able to communicate to different targets, customising tools and information systems. Communication services should aim at users, through systems directed at policymakers, stakeholders, media, civil society and other data users. These include web-based integrated tools like aggregator web-sites, wiki-like services, interactive briefing boards, multimedia newsletter and data journalism products.

The paper is presented for discussion to the Conference of European Statisticians seminar on "What is the value of official statistics and how do we communicate that value?"

* This document was submitted late due to the late submission of paper from the National Institute of Statistics.

GE.14-20573

1420573

Please recycle

I. Introduction

1. National Statistical Institutes (NSIs) are increasingly faced by new challenges both in terms of modernisation of statistical production processes and of the growing and increasingly differentiated information needs. The transition of official statistics from a niche service to a commodity able to guide individual, social and political behaviours, to inform the decision-making process, to help evaluate policies and foresee the consequences of today's choices, clearly highlights the role of NSIs as fundamental actors in the knowledge society. It also stresses the need to identify efficient and shared communication strategies to inform stakeholders.

2. In recent years, societies have been evolving towards a "Web 3.0 model", characterised by the on-going trends in web technologies and progress in information production, through the availability of big data and increasing demand for (linked) open data. Interoperability and availability of linkable accessible data are two key elements of this new trend. The consequences for users are that data supply has increased dramatically, but at the same time it becomes more difficult to make sense of information and use it to make decisions. The challenge for NSIs is to provide services that allow users to extract tailored information from data in a friendly and interactive way.

3. NSIs' response to these challenges has to focus on communicating and disseminating statistical information based on an integrated approach. This is built on the extensive use of new media, service-oriented websites and other on-line information applications to put users in the "driver's seat". New communication channels are required in order to reach a higher number of users, responding to specific and fast-changing demands. Such synergic approach can bring positive effects on "consumers" of statistical information, strengthening the perception of the innovative value added brought about by official statistics: the transformation of statistical data into knowledge. This can contribute to establishing and maintaining an active and collaborative dialogue with users (e.g., engagement, customer fidelity, sharing etc.), strengthening the reputation of NSIs.

4. The rest of this note will briefly touch on the innovation scenario for NSIs and consequences for dissemination and communication, as well as the transformation of traditional statistical outputs into knowledge-intensive user-centric services. It will present Istat modernisation strategy and the communication approach consistent with the innovation agenda. The note concludes with future developments and challenges.

II. The innovation scenario

5. NSIs have increasingly focused on revamping statistical production methods and products to meet new information needs. Priority was given to optimising statistical information production systems (cost-efficiency actions) and improving the quality and quantity of information products tailored to stakeholders needs (cost-effectiveness).

6. Modernisation strategy objectives that are shared across NSIs include (Figure 1) reducing the respondents' burden for households and companies and reducing the cost of the statistical production process. This is also combined with a trend towards general government modernisation reforms fostered by IT developments including the support to paperless processes. At the same time IT developments (including the development of open platforms and software built on the concept of interoperability) increase the availability of tools that allow users to generate interactively the information needed when they have to use it. This contributes to spurring evidence-based decision making by different stakeholders. The use of innovative dissemination tools can help better respond to fast growing user needs.

Figure 1
Modernisation Objectives

NSIs share common innovation goals
Ensuring cost-efficiency
Reusing services and infrastructure
Enhancing data accessibility and re-usability
Using innovative and effective communication tools
Responding to increasing and changing user needs
Supporting paperless government
Investing on innovation culture

Source: Istat, 2013

7. In accordance with these objectives, NSI modernisation strategies have been developed and implemented in response to these expanding needs. The change was focused on the standardisation and industrialisation of the statistical business chain, the implementation of re-usable infrastructures to support National Statistical Systems and NSI Networks, the adoption of a metadata-driven systems and the development of client-oriented tools to enhance the value creation in official statistics.

8. The implications for NSIs of this new environment cover several dimensions. These include the following need-based approach as a guidance to the design of communication strategies:

(a) listening and comparing different point of views (internally and externally) and utilising feedback to drive decisions;

(b) promoting and sharing innovation through simple messages and tools;

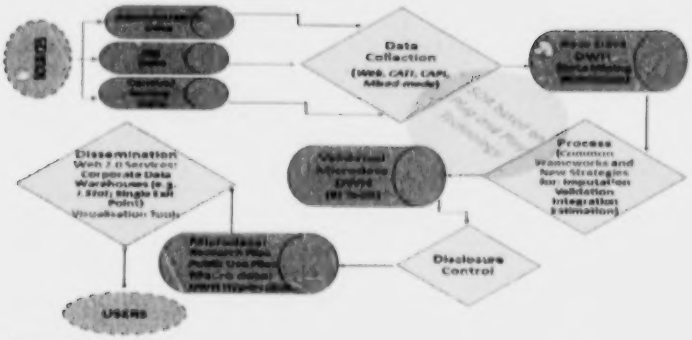
(c) joining forces to operate the selected innovation model;

(d) encouraging a cooperative environment to foster innovation (resources, exchange of ideas, common goals).

9. *Stat2015* is the flagship innovation programme that Istat has been carrying out since 2010: it is designed to allow the transition to a business environment moving from a stovepipe model to a new structure based on a Service-Oriented Architecture (SOA) supported by plug-and-play technology, consistent with the industrialisation and standardisation of production processes achieved by reusing information that has been already collected on respondents through a variety of sources and building platforms that allow seamless exchange and data integration.

10. The adoption of an Enterprise Architecture (EA) and of a Business Architecture (BA) is required to redesign statistical processes and to implement generalised services, methods and standards inspired by a common vision shared with the European and international statistical communities (Figure 2). This programme is both ambitious and challenging, requiring structural changes in the working environment, in its technological and methodological foundations.

Figure 2
Stylised Stat2015 framework



Source: Istat, 2012

11. *Stat2015* infrastructure is compliant with the SOA framework and aims at:

(a) collecting survey data, accessing administrative and other big data sources using digital technologies (e.g., from web portals, to computer-assisted interview collection, to administrative files, single entry point and web scraping tools);

(b) organising data in a raw data repository to be processed;

(c) carrying out data integration, validation and estimation tasks with generalised tools (e.g., methodological software, etc.);

(d) storing validated micro data in a data warehouse to be used for the following dissemination stage;

(e) disseminating micro and macro data through different tools reflecting user needs (e.g., from Micro data Files for Researchers through Research Data Centres and anonymous Public Use Files to the corporate macro data warehouse I.stat and related visualisation and communication tools).

12. Development projects included in the *Stat2015* programme are led by teams that involve human resources spanning across various Istat organisation units and focusing on specific stages of the GSBPM (Generic Statistical Business Process Model)-based production process. Specific resources in the budget are allocated to the programme projects and investment planning (e.g., IT and methodological support) reflects the high priority assigned to the programme. More in detail, for 2013 about one fifth of the whole budget has been allocated to *Stat2015*.

13. In the short run, *Stat2015* development is planned with an emphasis on the following key components of the programme:

(a) design and implementation of the shared Business Architecture (BA);

(b) design and implementation of the Corporate Metadata System (CMS);

(c) tools to enhance user need assessment capacity;

(d) data collection and management web portals;

(e) generalised methods and tools for data processing and analysis;

(f) new dissemination and communication tools for micro and macro data to enhance user interactions.

14. With regard to BA, this is central for defining, managing and monitoring the necessary changes for *Stat2015* development: this approach can guide the operational changes in the production chain, highlighting the required infrastructure that should be put in place under this framework .

15. CMS planned design and implementation are expected to consolidate existing systems (e.g., the quality metadata information system). The medium-term objective of the process is to adopt an Istat metadata architecture that guides users through all stages of the production chain.

16. Assessing user needs represents the very first stage for tailoring products and services, developing new tools and partnerships that allow Istat to assess what users really need. In this area, the following two topics feature prominently in the agenda of *Stat2015*:

(a) development of web-based tools to gather customer feedback to allow maximum user interactions with Istat. The key portal to collect these feedback and suggestions is the corporate web site www.istat.it, but also social platforms and networks are increasingly exploited to gather feedback in a dynamic way;

(b) redesign of the web portal to provide adequate information and specific tools, helping to disseminate the results of the innovation programme, its infrastructure and services among this network of public and private producers of statistics.

III. Istat "Communication-4-Innovation" strategy

17. Istat "communication-4-innovation" strategy is embedded in the *Stat2015* programme. Key communication actions are defined depending on the target. Specific user groups selected by Istat include, selected population groups (e.g., young, elderly, women), universities, socio-economic research bodies, decision-makers, opinion leaders, stakeholders, enterprises, media. The strategy also covers national and international institutions, general government and territorial bodies, international statistical bodies and institutions.

18. Key actions in this area include the following:

(a) profiling users to offer customised services/products through dedicated channels and make the information produced more accessible and usable;

(b) leveraging multimedia and digitalisation effectively by using a targeted multichannel strategy to offer products and services, especially through new digital communication channels;

(c) promoting and sharing innovation through simple messages and tools, maximising communication flows both inside and outside the Institute, thus fostering impartial access to information of all involved parties;

(d) monitoring the communication process through measurement systems of communication initiatives and actions, and the definition of tools and methods for research and analysis of communication impacts;

(e) creating international networks and participating in the most advanced communities of the sector.

19. More in general, to communicate the leading role in innovation, NSIs have to face the big challenge of adapting to the evolving internet world, mobile technologies and social media and then structuring their services and products accordingly. Istat strategy is based on promoting initiatives such as:

(a) sharing innovation-related best practices (also in the light of the feedback collected through reputation surveys), and the role of official statistics in policy assessment;

(b) endorsing the image of digitalised official statistics, by promoting technological innovation and efficient management and administration;

(c) fostering partnerships with leading operators and actors in the field of innovation on cutting edge issues (e.g., Google on the use of big data);

(d) adopting innovative communication standards: common terminology responding to an innovation-oriented community, shared graphical standards, enhancing data accessibility and re-usability;

(e) encouraging a policy of social responsibility to reduce environmental impact and respect international eco-friendly measures.

20. NSIs challenge is to be able to service this wide range of customer requirements taking into account different degrees of statistical literacy (Figure 3) through:

(a) customer satisfaction and customer profiling surveys like VoC (Voice of the Customer), aiming at identifying statistical information users;

(b) web analytics and systematic analysis of users' feedback and requests;

(c) "Guidelines on social media policy" to outline the communication strategy for the different social platforms (e.g. behaviours in directly managed social media pages, communication languages adopted, management and interaction modes etc.), with special attention dedicated to emerging social networks;

(d) responsive websites for mobile devices;

(e) data visualisation tools: info-graphic to communicate statistics from a narrative point of view besides complex data analysis and visual navigation tools;

(f) virtual press office, e.g. Web 2.0 press office;

(g) multi-functional and technological blogs to host and support informal meetings with different users;

(h) intranet portals able to supply adequate working tools and favour knowledge sharing to foster innovation (exchange of ideas, common goals).

Figure 3
Providing tools for different users

For Internal Users

- Intranet portal
- Sharing platforms
- Network blogs
- Virtual forums
- Web meetings
- Web seminars
- E-learning systems
- Seminars
- Workshops
- Training Courses

For External Users

- Aggregator web-site
- Web-based integrated tools
- Wiki-like services
- Interactive briefing boards
- Multimedia newsletter
- Data journalism

For Institutional Users

- Integrated information systems
- Dashboards for benchmarking
- Subject briefs
- Policy impact tools

21. In order to enhance data accessibility and re-usability, statistical information and analysis should be disseminated and communicated through an open strategy of data sharing. This can support the knowledge of economic, social and environmental trends and enhance the decision-making processes. Moreover, a coordinated development of integrated

information systems is needed to help access information and reduce statistical burden (e.g., by increasing the use of big data for communication purposes).

22. The availability of tools and services enhancing data access and re-use is fundamental to communicate innovation. In particular, priorities have been identified by Istat for the medium term in the development of the following:

(a) usable, responsive and accessible web-based integrated tools suitable for mobile devices, which can support interactive contents. These tools should be compliant with web-design structured taxonomies and semantics, taking into account web analytics and user research on search engines;

(b) web and API services based on shared standards and open formats;

(c) data visualisation and info-graphic dissemination tools;

(d) digital editing consistent with content, format and production process innovation, to be released on e-stores;

(e) apps and widget for mobile phones;

(f) web services to access micro data to favour the availability of public-use files;

(g) multimedia to disseminate statistical information, such as videos on key information areas based on integrated statistical data;

(h) endorsing of official statistics open production as main supplier of re-usable data also through social platforms.

23. Moreover, in order to orient communication and dissemination strategies to innovation, statistical organisations can share ideas, projects and also re-use tools and components in a plug-and-play fashion. Cooperation should also promote a common strategy on emerging issues such as open data and big data through:

(a) active participation in international and innovative networks;

(b) promotion of frequent informal meetings to exchange ideas and share innovative projects;

(c) use of common innovative tools developed through sectorial communities and social platforms (e.g., Facebook, YouTube, Flickr, SlideShare, Twitter).

IV. Future developments

24. As part of NSI innovation strategies, statistical data need to be proactively communicated by NSIs to engage users through a customer-centric approach being aware of user needs. Innovative communication tools can help respond to increasing and changing user needs. In particular, a clear and simple language, the correct use of social media, user-friendly web tools to browse and visualise data, the focus on re-use and open data are drivers to spread innovation in the statistical business chain, focusing on information demand.

25. The overall communication strategy of innovation is therefore based on the value provided by accompanying and enriching statistical information with services and tools easing its use and favouring its conversion into knowledge. In order to facilitate this transition, NSIs will have to adopt a balanced mix of new communication tools, so that the original message can remain substantially correct and consistent with the disseminating institution intention also in the subsequent web passages, e.g. when disseminated through social platforms.

26. This means that communicative efficacy will depend on high-quality, clear and precise messages, the only able to protect it from semantic distortions. The new client-oriented communication model will be enriched through the development of a collaborative, horizontal strategy based on listening, comparing different point of views and fostering the exchange of ideas. Within this framework, the new digital communication channels become a fundamental tool.

27. NSIs need to be able to communicate to different targets, customising tools and information systems. Communication services should aim at users, through systems directed at policymakers, stakeholders, media, civil society, and data users. These include web-based integrated tools like aggregator web-sites, wiki-like services, interactive briefing boards, multimedia newsletter and data journalism products. Communication targets also include NSI staff who should have access to tools like intranet portals, sharing platforms, web seminars, and e-learning systems.

V. References

Baldacci, E. (2013). Innovate or perish – Italy's *Stat2015* modernisation programme. 59th ISI-World Statistics Congress, Hong Kong, 25-30 August.

http://www.statistics.gov.hk/wsc/IPS085-P3-S.pdf.

Barcaroli, G., Falorsi, P.D., Fasano, A., Mignolli, N. (2013). Business Architecture Principles to Foster Industrialisation and Standardisation at the Italian National Institute of Statistics. 59th ISI-World Statistics Congress, Hong Kong, 25-30 August.

http://www.statistics.gov.hk/wsc/STS048-P4-S.pdf.

Statistical Network *Innovation in dissemination* (SNID, 2013). http://www1.unece.org/stat/platform/display/statnet/Innovation+in+Dissemination

Statistical Network *Business Architecture Project* (2013). http://www1.unece.org/stat/platform/display/statnet/Business+Architecture+Project

UNECE/Eurostat/OECD (2013). *Generic Statistical Business Process Model (GSBPM version 5.0)*. http://www1.unece.org/stat/platform/display/GSBPM/Generic+Statistical+Business+Process+Model

UNECE (2013). Work Session on the Communication of Statistics. http://www.unece.org/index.php?id=3214

This document was submitted late due to the late submission of paper from the National Institute of Statistics.

United Nations ECE/CES/2014/26

Economic and Social Council

Distr.: General
29 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration Statistics

Measuring integration of immigrants and descendants – Danish experiences with register-based data

Prepared by Statistics Denmark

Summary

The paper describes the advantages of a register based system with complete data in conditions where populations are diverse regarding age composition. The paper introduces Statistics Denmark's definition of immigrants, descendants and persons of Danish origin and explains how this approach is preferable to a definition based solely on the place of birth or citizenship. Some examples of labor market and education statistics are also given where it is necessary to consider immigrants' and descendant's characteristics regarding age composition and educational preferences. The paper notes the importance of data on educational attainment as the Danish register based system is unable to provide sufficient information about immigrants' educational attainment.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

GE.14-20522

1420522

Please recycle

I. Definition of ancestry

1. Denmark has had a register-based population register since 1968. Every person in the Danish population has a unique identification number (CPR), which is widely used in issues related to social security, health care, pensions, education, bank accounts, etc. Therefore, data can be linked using the CPR which makes it possible not only to have a total count of the population broken down by age, sex, etc., but also complete statistics covering the entire population about socioeconomic status, educational attainment, etc.

2. In the beginning of the 1990s, political interest emerged to identify immigrants and descendants in statistics. Statistics Denmark decided to create a new variable called 'ancestry', where each individual in the population was placed in one of the following three categories: 1) Danish origin, 2) immigrants and 3) descendants.

3. The variable was made using already available information in the population register about place of birth, citizenship, the parents' place of birth and the parents' citizenship.

(a) A person is categorized as 'Danish origin' if he/she has at least one parent who is both born in Denmark and has Danish citizenship. The person does not have to be born in Denmark in order to be categorized as Danish origin;

(b) A person is categorized as an immigrant if he/she is born abroad and has no parent who is both born in Denmark and has Danish citizenship;

(c) A person is categorized as a descendant if he/she is born in Denmark and has no parent who is both born in Denmark and has Danish citizenship.

4. If a person´s parents have never been registered in the Danish population and no information is available about them, the person´s own birthplace and citizenship decides ancestry. A person born in Denmark would be categorized as a descendant if he/she has foreign citizenship and Danish origin if he/she has Danish citizenship. A person born abroad would be categorized as an immigrant.

5. Immigrants and descendants can also be broken down by country of origin. When both parents are known the country of origin is categorized as the mother´s country of birth. When only one parent is known the known parent´s country of birth will decide the country of origin.

6. When no parents are known the person´s own country of birth is defined as country of origin for immigrants. For descendants, who are all born in Denmark, the person´s country of origin is defined as the person´s citizenship country.

7. Statistics Denmark often divides all different countries of origin into two groups: Western countries and non-Western countries. Western countries are defined as all European Union countries plus Andorra, Iceland, Liechtenstein, Monaco, Norway, San Marino, Switzerland, the Vatican, Canada, United States, Australia and New Zealand. All other countries are considered non-Western in this context.

Table 1

Population in Denmark broken down by ancestry and birthplace, 1 January 2013

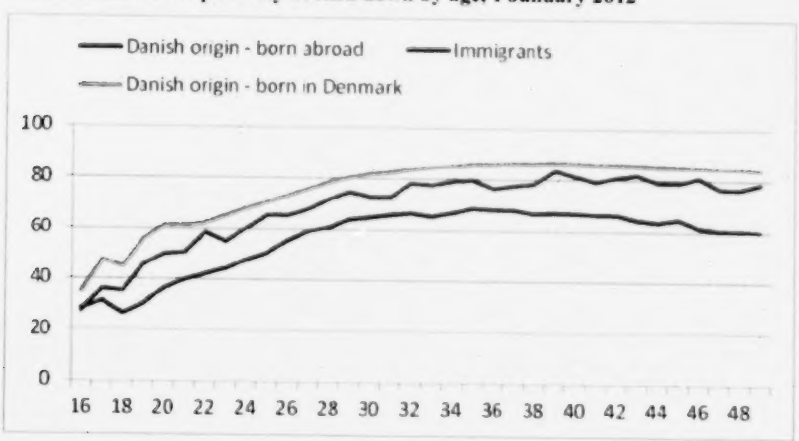
	By ancestry	*Born in Denmark*	*Born abroad*	*Percentage born in Denmark*
Total	5 602 628	5 080 404	522 224	90.7
Danish origin	5 001 954	4 936 116	65 838	98.7
Immigrants	456 386	0	456 386	0.0
Descendants	144 288	144 288	0	100.0

8. A very important aspect of the Danish definition is that it allows persons born abroad to be categorized as Danish origin. As can be seen from the table above 65.838 persons with Danish origin were actually born abroad. Compared to the total group of persons with Danish origin it is a very small group which only constitutes 1.3 per cent of the total. But if you look at all the persons born abroad in the Danish population their share is 12.6 per cent.

9. An important argument for using ancestry as a variable, instead of just dividing the population as born abroad and born in Denmark, is that if a person is born abroad and has a parent who is both born in Denmark and has a Danish citizenship, the person has more in common with other persons with a similar background born in Denmark than with other people born abroad. In most cases persons with Danish origin born abroad, who become a part of the Danish population, immigrate to Denmark while they are still children.

Figure 1

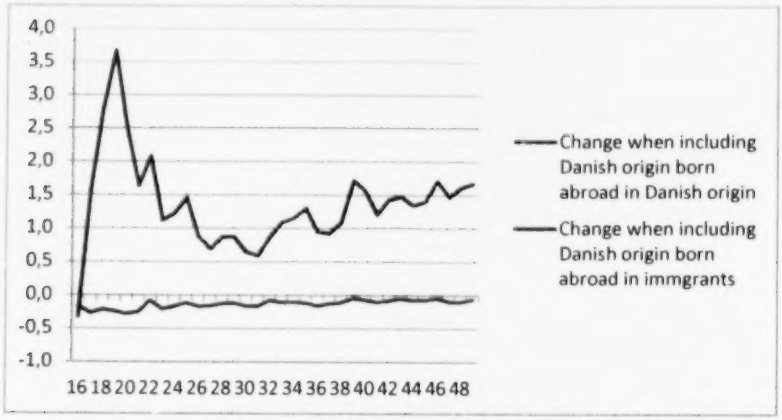
Employment rates for male immigrants and men with Danish origin born in Denmark and born abroad respectively broken down by age, 1 January 2012



10. The figure above shows that as far as employment rates are concerned persons of Danish origin are closer to persons with Danish origin born in Denmark than they are to immigrants. Even though employment rates are lower for persons with Danish origin born abroad than persons with Danish origin born in Denmark, it is preferable to have them in the same category because their influence is negligible on the complete category of persons with Danish origin.

Figure 2

Percentage point change in employment rates for male immigrants and males with Danish origin when including Danish origin born abroad broken down by age, 1 January 2012



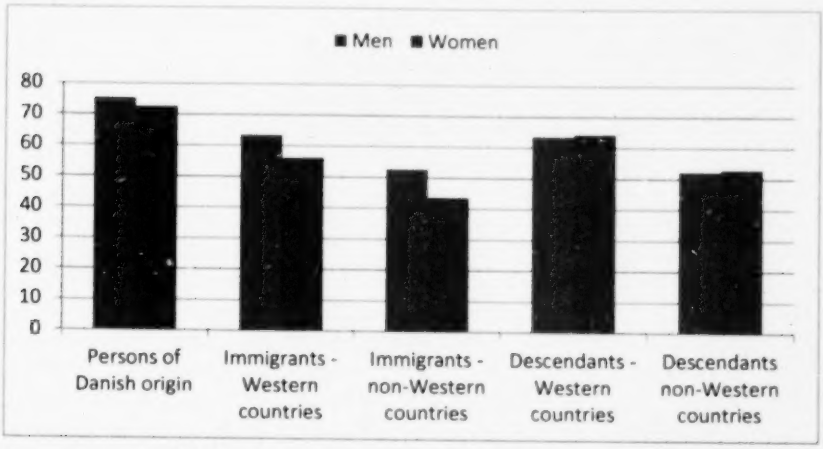
11. Including Danish origin born abroad as immigrants would overestimate employment rates for immigrants at some age levels by more than 3.5 percentage points with an average of 1.36 for all the age levels from 16 to 49. On the other hand, including Danish origin born abroad in Danish origin only causes an average negative change in the employment rate by 0.13 percentage points. Danish origin born abroad is not a very interesting group by itself in terms of integration, while immigrants are. So, clearly it is preferable to present statistics on Danish born abroad separately from other persons born abroad by including them in Danish origin.

II. Employment

12. Employment rates are often considered one of the most important indicators for integration. A simple indicator is the employment rate for 16-64-year olds broken down by gender and ancestry.

Figure 3

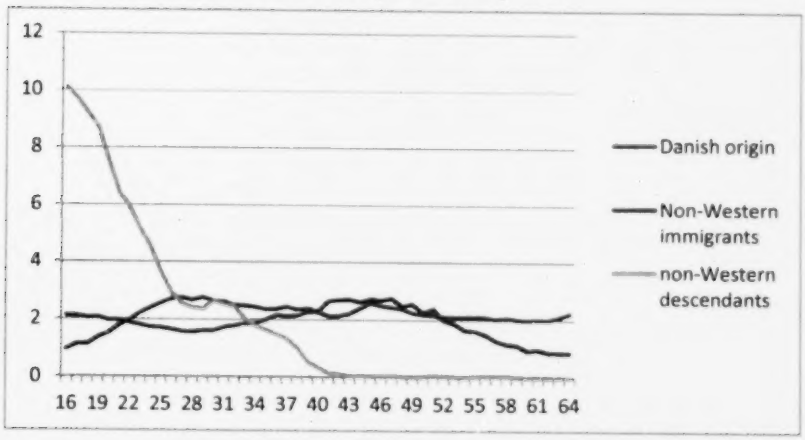
Employment rates for 16-64 year olds broken down by gender and ancestry, 1 January 2012



13. One noticeable result from the figure above is that non-Western male descendants have a slightly lower employment rate than non-Western male immigrants when analysing all 16-64 year olds. A very surprising result when you consider that the descendants are born in Denmark. However, it is also a result that does not take the huge age variations between the different groups into account.

Figure 4

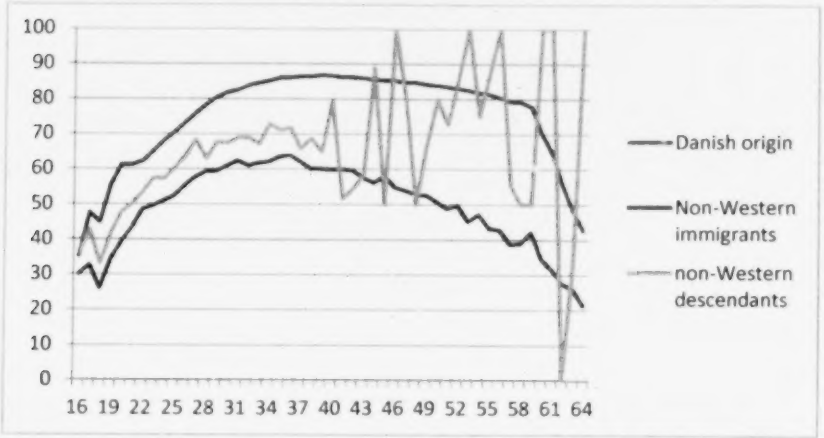
Age level's percentage of all 16-64 year old for non-Western male immigrants and descendants and males with Danish origin, 1 January 2012



14. Non-Western descendants are very young and have an age distribution that is strikingly different from both non-Western immigrants and males with Danish origin. Non-Western descendants are hugely overrepresented in the lower age levels with 52 per cent of all 16-64-year old in the first six age levels alone (16-21). For non-Western immigrants 16-21 year olds only accounts for 7 per cent of all 16-64 year olds. On the other hand the population of non-Western descendants above the 40 years level is practically non-existent in Denmark.

Figure 5

Employment rates for non-Western male immigrants and descendants and males with Danish origin broken down by age, 1 January 2012



15. At the same time employment rates vary significantly with age as can be seen from the figure above. They are particularly low for the group of young people where the overrepresentation of non-Western descendants is largest and they are the highest for persons in the 30´s and 40´s where the population of non-Western descendants is very small. In the figure the employment rate for every age level between 16-64 years is shown, but for persons above 40 they are based on very small populations for non-Western descendants, which is the reason why they are so volatile.

16. Comparing employment rates for non-Western immigrants and descendants by age level, it is obvious that descendants have higher employment rates for almost all the age levels. The few age levels where this is not the case have a very small population of non-Western descendants and therefore have no influence on the overall employment of 16-64 year olds. However the overall employment rate is still somewhat higher for non-Western immigrants (52.7) than non-Western descendants (52.2). It can seem like a paradoxical result, but it is of course perfectly possible given the completely different age distribution of the two populations.

17. Therefore, the overall employment rate for all 16-64 year old can be a misleading indicator for descendants. If the need is for one simple number it would clearly be a better idea to look at for example 30 year olds. In this case one would have an employment rate for non-Western male descendants of 68 per cent which can be compared with 61 per cent for non-Western male immigrants and 82 per cent for males with Danish origin (82 percent). A huge benefit of having a register-based system is that you can work with the total population at all age levels. Therefore, most of the time calculations for individual age levels are reliable for all groups, and are often preferable when comparing populations with very different age distributions.

III. Educational attainment

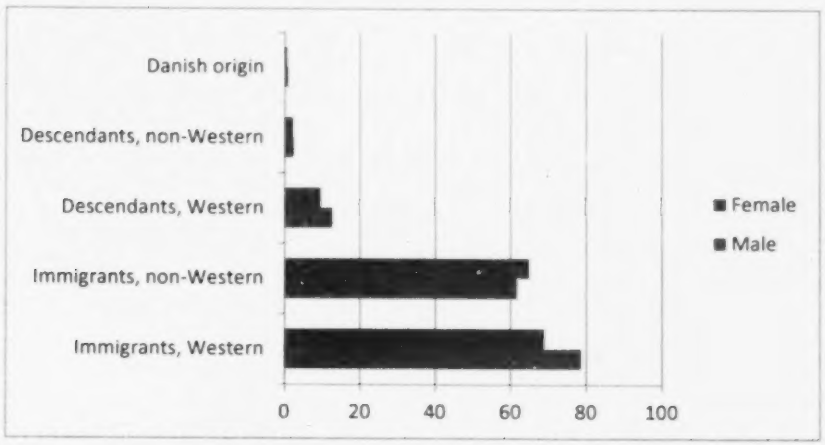
18. The foundation for Danish statistics on education is data from Danish education institutions at all levels. Records in these data sources are created when a person starts in a particular education, when he/she finishes the studies or if the person abandons the studies without finishing them. In many ways it is a unique system which has almost endless

possibilities in terms of analyses, as persons can be identified and linked with other relevant data – for example labour market or social benefits data.

19. However, when it comes to matters related to integration the system is of limited value. In practice only education completed in Denmark is registered. Most immigrants have no completed Danish education, as most have immigrated to Denmark as adults.

Figure 6

Percentage with no Danish education by 16-64 year olds broken down by ancestry. 1 January 2013

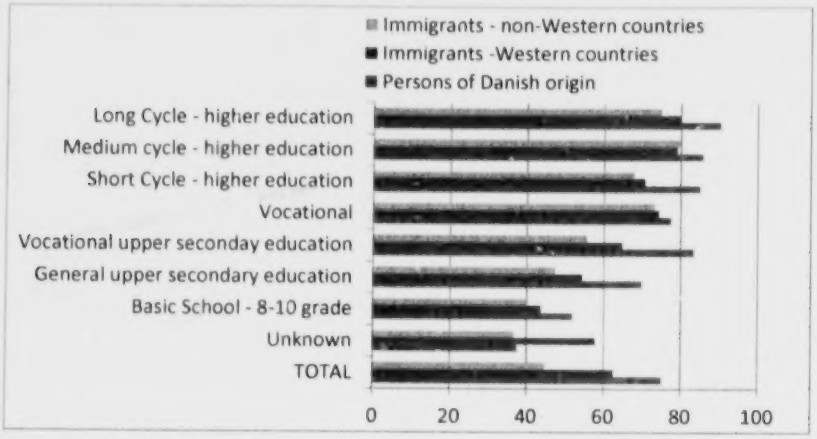


20. Over the years, some efforts have been undertaken to fill out missing information for immigrants by conducting surveys. However, response rates have been poor and the answers received are not completely comparable with data collected from Danish education institutions. There have been attempts to use survey data as inputs for an imputation model. But in general, the problem with lack of data is so significant that no solution, however advanced it may be, will be able to produce data on educational attainment for immigrants which can be compared with other groups in a meaningful way. Furthermore, a solution based on imputation also makes it difficult to use data at the most detailed level.

21. Analyses and indicators on education and integration should therefore be restricted to education completed in Denmark, which rules out most of the immigrant population. However, it is still possible to compare persons who have actually completed a given education in Denmark with persons of Danish origin with the same educational background. Linking education data with labor market data, one could for example calculate employment rates for these particular groups. It is relevant in order to measure whether immigrants with Danish education at a given level have similar employment rates to persons of Danish origin with the same education level. One should also keep in mind that this group of immigrants constitutes a relatively low proportion of all immigrants. In addition, immigrants with Danish education have overall been in Denmark for a longer time than other immigrants, which is something that alone could have an influence on employment.

Figure 7

Employment rates for 25-64 year old females broken down by ancestry and educational level. Only Danish education, 1 January 2012

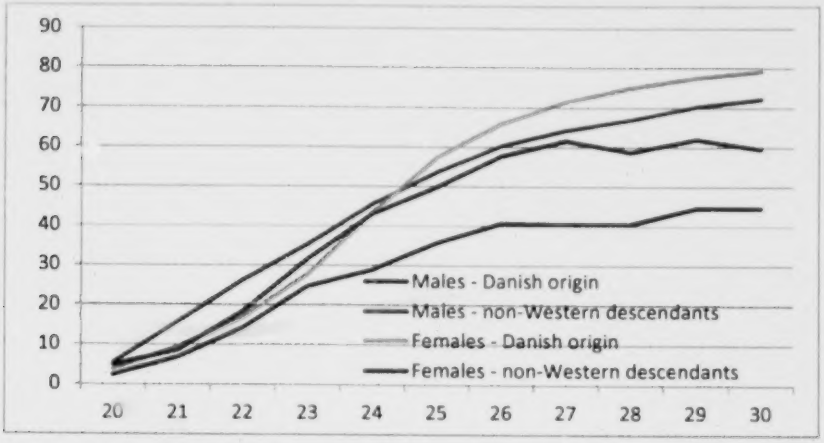


22. Immigrants do not have as high employment rates as persons with Danish origin, even when they have completed Danish education at the same level. However, differences are smaller when you look at a given educational level compared to the overall situation. For all 25-64 year old females only 45 per cent of the non-Western immigrants were in employment compared to 75 per cent for women with Danish origin. Only the levels Vocational and the three levels of higher education give qualifications which can be used directly in the labor market. Among these levels the differences are clearly smaller, especially for the levels Vocational and Medium cycle – higher education where non-Western immigrants only lag employment rates of women with Danish origin by 4 and 6 percentage points respectively. Looking specifically at the group of immigrants who have actually completed education in Denmark, is an example of how educational data can still be used for relevant statistics even though overall coverage for immigrants is very low.

23. Descendants, who are all born in Denmark, are easier to compare with persons with Danish origin in terms of educational attainment than immigrants. However the issue about their age distribution described in the labor market section is also highly relevant regarding education. Non-Western descendants have an age distribution where it is in practice only young people's educational attainment which can be analyzed. Regarding educational attainment among young people, it is in general a good idea to look at single age levels instead of age groups, because educational attainment changes rapidly at young age levels. The question is which age levels are preferred in order to have the best indicator for comparing descendants and persons with Danish origin.

Figure 8

Percentage with a completed Danish education at vocational level or above, broken down by age an ancestry, 1 January 2013



24. Vocational level or above comprises vocational, short cycle higher education, medium cycle higher education, bachelor and long cycle higher education. From the figure it can be seen that the percentage of persons with a completed education at these levels changes fast in the 20s. It is not until the late 20s that some stability is achieved.

25. In a Danish context it would clearly be misleading as an integration indicator to look solely at the percentage with a completed education at vocational level or above among 24 year olds. Among women one would find the percentage is 43 among both women with Danish origin and non-Western descendants. This could lead to the conclusion that integration is perfect and that non-Western descendants have the same educational attainment as women with Danish origin.

26. Looking at what happens in the figure after the age of 24, one can see that this would clearly be a wrong conclusion. The percentage continues to rise for women with Danish origin, while it stagnates for non-Western descendants. Among 30 year olds, 79 per cent of the women with Danish origin have completed an education at vocational level or above, while the percentage is only 60 among non-Western descendants – a difference of 19 percentage points.

27. The reason for different results depending on age level is that non-Western descendants, to a higher degree, complete education at the vocational level which they finish early. While women with Danish origin, to a higher degree, complete higher education cycles, which in many cases they do not finish until they are in their late 20s

28. Preferences regarding education are different among persons with Danish origin and non-Western descendants. Therefore, if the idea is to have an indicator with the intention of measuring educational attainment among young people, one should pick a single age level where education is more or less complete. In the Danish context, 30 years seem like a good choice.

United Nations ECE/CES/2014/27

Economic and Social Council

Distr.: General
29 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration statistics

Migration statistics in Switzerland: New sources and challenges

Prepared by the Swiss Federal Statistical Office

Summary

This paper discusses a number of issues related to the availability of new migration data sources in Switzerland and the collection of migration core variables through the population and housing census. The paper examines the meaningfulness of selected census variables for migration analyses and compares, for example, the census core variable 'place of usual residence one year prior to the census' with annual migration flow data. Several data sources allow the identification of population groups relevant to international migration, for instance population registers, census sample survey, and the labour force survey. This paper considers whether these different data sources yield the same results and how to deal with conflicting data.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

GE.14-20510

1420510

Please recycle

I. Introduction

1. Migratory movements which are generally increasing in both complexity and volume constitute an important factor of demographic and economic growth in many countries. International migration is a topic that constantly draws attention from policy makers, the media, the research community as well as the general public. Comprehensive and reliable statistics on migration are therefore required to address these user needs.

2. The flows and stocks of migrants are the two more frequently used concepts in measuring migration. Migration flows represent the movements of persons to or from a country or a territorial division during a given period of time. The migrant stock identifies the part of the population with a migration background staying in a country or a territorial division at a given moment of time. Measures of international migration are typically derived either from census and survey data or from data from administrative sources.

3. This paper discusses a number of issues related to the availability of new migration data sources in Switzerland and the collection of migration core variables as outlined in the *Conference of European Statisticians (CES) Recommendations for the 2010 censuses of population and housing*. The main objectives are twofold: to present some new findings as well as still existing gaps and challenges in Swiss migration statistics and to provide a critical appraisal of the intended purpose of two of the suggested core migration variables.

4. The section II of the paper briefly describes the new register-based census in Switzerland and its contribution to the collection of migration data. The section III explores three aspects of international migration statistics: (1) the number and characteristics of ever-international migrants living in the country, (2) the analysis of the mobility question about the place of residence one year ago, and more specifically, a comparison of annual migration inflows derived from stock data and register-based flow data, (3) the identification of population groups relevant to international migration according to different data sources. Finally, some general conclusions are presented in the last section of the paper.

II. The new Swiss census system

5. In Switzerland, the traditional decennial census – based on a full field enumeration, with information collected using census forms – has been replaced by an integrated statistical system which provides annual data. The new census system combines the use of administrative registers with sample surveys. The move to annual data availability allows for regular monitoring of the population structure and its changes. The first reference day for the new census was on 31 December 2010.

6. The new system consists of four different annual surveys, among them – and of particular importance to migration statistics – the register survey and the new large-scale structural survey. The register survey exploits existing and harmonized administrative data from both centralized and local population registers: (1) alien registers maintained by the Federal Government covering non-nationals living in the country and (2) municipal or cantonal registers of the resident population. Additionally, data from the Federal Register of Buildings and Dwellings enable analyses at a very detailed geographical level. The structural survey is an annual sample survey of 200,000 people. It includes variables traditionally collected by the census but not currently available in registers, thus complements the register survey. It covers people living in private households who are aged 15 or over and involves the self-completion of paper questionnaires or on the internet by respondents.

7. The social security number which uniquely identifies a person plays an important role both in the maintenance of administrative registers and in their statistical use, particularly in linking information from register and survey data[1].

8. The new census system provides previously unavailable information about the stock of immigrants and other persons with a foreign background, and also on the timing and the geographical patterns of their migration. It is useful for capturing migration since data are available on an annual basis and not only at intervals of five or ten years, as in a traditional census. Due to its primary reliance on administrative data it is particularly well suited for producing data on migration flows: Registers are dependent on the timely registration of the population and keep track of all changes of residence, thus ensuring a continuous update of individual records. For this paper, data from the first three years of the new combined census (2010-2012) were analyzed.

III. Three selected issues related to Swiss migration statistics

9. This chapter discusses three aspects relevant to migration statistics. It focuses on three core variables of the CES 2010 census recommendations[2], explains how this information is collected by the Swiss Federal Statistical Office (FSO), presents the most recent findings and addresses some major challenges linked to the available data and the purported relevance of the core migration variables.

A. Ever-international migrants

10. The *Recommendations on Statistics of International Migration*[3] define an international migrant as 'any person who changes his or her country of usual residence'. In order to identify and quantify the stock of ever-international migrants in a country, i.e. the set of persons who have ever changed their country of usual residence, the CES recommendations include the core topic 'ever resided abroad and year of arrival in the country' (§§ 379-381).

11. To collect information on this topic, individuals should be asked whether they have ever had a usual residence abroad. For those who ever resided abroad, the year of last arrival in the current country of residence, should also be collected. It should be noted that this variable does not provide any information on the previous place of usual residence (and year of arrival in the current place). Moreover, it cannot identify people who have established residence in the country on more than one occasion.

12. In the annual Swiss structural survey the following question is asked: 'Have you ever lived for at least one year continuously in a country other than Switzerland?' Respondents can reply either no or yes. In the affirmative case, the year of last arrival (or return) to Switzerland must be indicated.

13. According to the 2010 and 2011 structural surveys about one fifth of all permanent residents of Switzerland are ever-international migrants. (2010: 22.0 per cent; 2011: 18.9 per cent). This corresponds to about 1.2 to 1.4 million individuals. However, as up to half

[1] The Federal Statistics Act of 1992 (art. 14a) sets the legal framework for any data linkages carried out by FSO.

[2] UNECE, Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing, United Nations, New York and Geneva, 2006 (Chapter VIII. International and Internal Migration, pp. 83-93)

[3] UNSD, Recommendations on Statistics of International Migration – Revision 1, Statistical Papers, Series M, No. 58, Rev. 1, United Nations, New York, 1998 (§32)

of the foreign-born declared – somewhat surprisingly – to have never resided abroad for at least one year, this percentage is definitely too low. According to the CES recommendations the group of ever-international migrants includes all foreign-born as it is assumed that they all resided or were expected to reside in the country of birth for at least one year (§365). Obviously, a significant number of respondents did not understand this question correctly. The share of foreign-born who answered illogically even increased from 36.7 per cent in 2010 to 50.6 per cent in 2011. Adjusting the data by assigning a migration background to all foreign-born, the share of ever-international migrants currently living in Switzerland is estimated to be about 34 per cent[4].

Table 1

Permanent resident population who ever resided abroad by country of birth, 2010 and 2011

Country of birth	*2010*		*2011*	
	in 1000	*in %**	*in 1000*	*in %**
Total	**1436.1**	**22.0**	**1244.4**	**18.9**
Switzerland	362.1	7.8	356.6	7.7
abroad	1053.1	57.2	879.3	45.6
unknown	21.0	39.7	8.5	39.0

*of total population of each 'country of birth' sub-group

14. A more detailed analysis reveals that 71 per cent of all ever-international migrants were born abroad. Nationals and foreign citizens are represented equally (49.5 per cent and 50.5 per cent, respectively). Whereas 42.9 per cent arrived in Switzerland after 1999, 38.2 per cent between 1970 and 1999 and 12.3 per cent before 1970, the breakdown by country of birth (Switzerland/abroad) discloses significant differences between the two groups. Among the native-born 29.1 per cent arrived in Switzerland after 1999; this proportion is considerably higher among the foreign-born (48.4 per cent). Moreover, foreign-born ever-migrants are younger (45.1 per cent were born after 1970 and only 18.0 per cent before 1950) than their native-born counterparts (with only 28.4 per cent born after 1970 and one third born before 1950).

[4] Correcting the collected data in such a way is relatively straightforward. However, the year of last arrival is missing for all reallocated individuals.

Table 2

Ever-international migrants by country of birth and year of last arrival and year of birth*, 2011

	Country of birth	
Year of last arrival	***Switzerland***	***abroad***
before 1950	0.9	0.9
between 1950 and 1959	4.1	2.7
between 1960 and 1969	10.3	7.5
between 1970 and 1979	13.5	6.8
between 1980 and 1989	15.5	11.5
between 1990 and 1999	18.3	16.3
2000 or later	29.1	48.4
unknown	8.3	5.9
Year of birth		
before 1950	33.9	18.0
between 1950 and 1959	18.3	14.2
between 1960 and 1969	19.4	22.6
between 1970 and 1979	14.9	23.8
between 1980 and 1989	9.3	16.6
between 1990 and 1999	4.3	4.7

*as % of all ever-international migrants

15. The figures above clearly show that the group of ever-international migrants is very heterogeneous. Broadly speaking, two large groups can be distinguished: 'life-time migrants' and 'return migrants'. A 'life-time migrant' is a person whose current place of residence is different from his/her place of birth[5]. A 'return migrant' is a person whose current place of residence is the same as the place of birth, but whose duration of residence is less than his/her age. Specifically, the latter group comprises persons without a foreign background (according to the definition in §398 of the CES recommendations) who, however, have experienced an international migration (e.g. due to having worked or studied abroad for at least one year).

16. For any in-depth analysis this item has to be cross-referenced with other socio-demographic variables. However, a basic disaggregation by sex, current age and citizenship is not nearly sufficient to provide meaningful insights. Additional variables have to be taken into account, some of which (such as country of birth, year of last arrival, age at the time of arrival) are most likely available in population censuses. Others (such as reason of immigration or return migration, total duration of residence abroad), however, are usually not collected by censuses and are – as is the case in Switzerland – also not available in administrative registers. Ideally, individuals' entire migration histories would be known if a thorough analysis and portrayal of ever-international migrants were to be attempted.

17. The CES recommendations point out that this topic will better capture international migrants living in a host country as not all migrants can be identified from information collected by country of birth or country of citizenship alone. However, the group of ever-international migrants is so diverse that other variables commonly available in censuses or

[5] This term is admittedly somewhat ambiguous as 'life-time migrants' may not live outside their country of birth for their entire lives.

in most administrative registers, will hardly be sufficient for a comprehensive analysis and thorough understanding of this group.

B. Migration stock data versus annual migration flow data

18. The migration core topic 'previous place of usual residence and date of arrival in the current place' (CES recommendations, §§ 384-389) provides information on geographic and time patterns of migration moves to the current place of residence. In operational terms, this topic can be implemented in either an extensive mode (by enquiring into year and month of arrival in the current place of usual residence and the previous place of usual residence) or a reduced mode (by asking about the place of usual residence one year prior to the census).

19. In Switzerland, the reduced mode is employed and the information is obtained from population registers. A snapshot view of internal and international migration flows can be derived from this variable as it allows for the identification of persons who have changed their usual place of residence at least once in the last calendar year. The topic is primarily intended to allow patterns of recent immigration to be studied, especially in countries that do not collect complete annual flow data.

Table 3

Permanent resident population by place of residence one year ago*, 2010-2012

Place of residence one year ago	*2010*	*2011*	*2012*
same municipality	92.5	92.3	92.3
other municipality in the same canton	3.6	3.6	3.5
other municipality in another canton	1.4	1.5	1.4
abroad	1.3	1.3	1.4
not yet born	1.0	1.0	1.0
unknown	0.1	0.4	0.3

*as % of total population

20. Results from the register surveys of 2010, 2011 and 2012 show the extent of geographic mobility in Switzerland. In any given year, 92 per cent of the population lived in the same municipality as one year before and can be broadly classified as 'non-movers' (even though an unknown number of them would have moved to another dwelling within the same municipality[6]). About 5 per cent of all individuals changed their place of residence within Switzerland and are internal migrants. The share of the population who had immigrated to Switzerland from abroad during the calendar year under consideration amounted to 1.4 per cent.

21. As far as the annual number of immigrants from abroad is concerned, register-based stock data – more precisely, the variable 'place of residence one year ago' – reveal that 104,000 and 115,000 individuals arrived in Switzerland in 2011 and 2012, respectively. A large majority of them (slightly above 60 per cent) were citizens of an EU-27 or EFTA member state, 16 per cent were Swiss nationals and about 12 per cent possessed the citizenship of a non-European country.

22. Because the register survey also allows obtaining annual flow data, there is the possibility to compare annual international inflows based on both stock and flows. Register-derived flow data put the annual number of immigrants at a somewhat higher level: about

[6] The municipality is the smallest civil division identified in official Swiss migration statistics which do not measure the change of the usual address.

150,000 per year. As annual migratory movements are depicted more accurately by flow data, the findings indicate that stock data based on the person's previous place of residence one year ago underestimate the annual international inflows by 43 per cent in 2011 and by about 30 per cent in 2012. A breakdown by citizenship of the immigrants further shows that the undercount is significantly higher for non-European citizens (57.7 per cent in 2011 and 46.0 per cent in 2012) than for citizens of an EU-27/EFTA member state (39.1 per cent and 25.6 per cent, respectively). The figures for Swiss nationals correspond more or less to those of the total of all immigrants. Consequently, immigrants from outside of Europe account for a higher share of all immigrants in the flow data (+2 percentage points).

Table 4

Annual immigration flows according to stock and flow data, by citizenship, 2011 and 2012

Citizenship	*2011*		*2012*	
	stock data	*flow data*	*stock data*	*flow data*
Total	**104,137**	**148,799**	**115,103**	**149,051**
Switzerland	16,488	24,104	18,241	24,006
EU27/EFTA	63,359	88,152	72,061	90,495
other Europe	7,618	10,247	8,132	10,219
other continents	16,668	26,281	16,659	24,317

23. The specific reasons for these discrepancies would have to be analyzed in greater detail. However, it is generally agreed that stock (or census) data have some well-known limitations as a source of inflow data and show only selective aspects of the migratory process as they are focused on migrants rather than migration events. Only immigrants that are still living in the country at the time of the census are counted, thus excluding those immigrants that have emigrated before the census date (e.g. in the same calendar year). Additionally, stock data are not able to capture multiple migrations of one individual that occurred between two subsequent census counts (e.g. the increasingly important phenomenon of circular or repeat migration), and emigration cannot be measured at all. Data on stocks of immigrants, therefore, rather represent the cumulative effect of net migration flows over a certain period in the past. While an acceptable proxy if there are no or only incomplete flow data, stock data are not an adequate substitute to estimate inflows if annual flow data are available.

C. Population groups relevant to international migration

24. The CES 2010 recommendations provide an excellent and comprehensive overview of the different population groups that are relevant to the capture and the analysis of international migration. The two derived non-core topics 'persons with foreign/national background' (§§ 398-402) and 'population groups relevant to international migration' (§§ 403-405) present some useful analytical classifications. They are based on the two core topics 'country of birth' and 'citizenship' and the non-core topic 'country of birth of parents'.

25. In Switzerland, the classification of the population stock by country of birth, (current) country of citizenship, citizenship at time of birth and country of birth of parents (and/or a combination of those four variables) is used for a national typology of the permanent resident population according to its migration status. Two major groups can be distinguished: (1) persons without a migration background (which broadly corresponds to the 'persons with national background' of the CES recommendations) and (2) persons with a migration background (or 'persons with foreign background') which can be further sub-divided into first- and second-generation immigrants.

26. This classification has been very useful for a more comprehensive monitoring and analysis of the impact of international migration because the basic distinction between foreigners and nationals is hardly sufficient anymore. Since the variable 'country of birth of parents' is not available in the population registers, survey data has to be used to identify the population with a migration background. However, household sample surveys conducted by FSO usually cover people who are aged 15 or over (and live in private households). Thus the precise migration status of persons below the age of 15 cannot be determined.

27. 0- to 14-year-olds account for 15 per cent of the permanent resident population of Switzerland according to the register survey of 2012. A breakdown by citizenship and country of birth further reveals that this proportion is slightly higher than the above average for the native-born (18.7 per cent) and for foreign citizens (15.9 per cent). Especially noteworthy is the fact that 56.8 per cent of all non-nationals born in Switzerland are less than 15 years old. A broad guesstimate based on the available register data puts the proportion of 0- to 14-year-olds with a migration background at about 27 per cent. However, it excludes naturalized Swiss citizens and includes native-born foreign children whose parents were also native-born. In conclusion, it can be argued that under 15 year-olds with a migration background constitute a sizeable population group which is generally excluded from sample household surveys and about which migration-related knowledge is insufficient.

28. Data from two different sources can be used for the classification of the population stock by migration status: (1) the structural survey and (2) the Swiss Labor Force Survey (SLFS). Both ask for the parents' countries of birth. While the SLFS also inquires about 'country of birth' and 'citizenship' of the respondents, these variables are added to the structural survey dataset from population registers via the unique person identifier.

29. According to the SLFS of 2012 64.5 per cent of the permanent resident population aged 15 and over have no migration background, while 34.7 per cent of all adults have a migration background. For 0.8 per cent of the population the migration status cannot be ascertained due to missing or implausible answers to questions used for its determination.

30. The structural survey of 2011 leads to slightly different results: 62.8 per cent of the population above the age of 14 has a migration background and 34.0 per cent of them do not. The proportion of adults for whom migration status cannot be determined stands at 3.2 per cent and is thus 2.4 percentage points higher than in the SLFS.

31. The two surveys reveal more important differences for specific population sub-groups and/or when broken down by sex and age. For example, the number of native-born Swiss with two foreign-born parents is 47 per cent and that of second-generation non-nationals[7] 27 per cent higher in the SLFS than in the structural survey. On the other hand, the structural survey leads to a higher number of foreign-born nationals (+11.1 per cent).

[7] native-born foreigners with at least one parent born abroad

Table 5

Permanent resident population (age 15 and over) by migration status and data source*, 2011 and 2012

Population Group	*Labor Force Survey 2012*	*Structural Survey 2011*
persons without a migration background	**64.5**	**62.8**
foreign-born	1.0	1.1
native-born	63.5	61.7
persons with a migration background	**34.7**	**34.0**
foreign-born	27.6	27.9
non-nationals	19.3	19.5
nationals	8.3	8.4
native-born	7.1	6.1
non-nationals	2.8	2.2
nationals	4.3	3.9
migration status unknown	**0.8**	**3.2**

*as % of total population

32. As the two reference periods are different (end of 2011 and 2nd quarter of 2012, respectively) some of the discrepancies can be explained by the influence of migratory events of the first half of 2012. Nevertheless, the results of the two surveys – although not contradictory – pose some challenges for data dissemination and interpretation. Which are the 'correct' data? Should there be only one reference source for dissemination purposes? Or should divergent data be made available to users of official statistics? If so, how can the reasons of such differences be explained in a transparent and comprehensible way?

33. Survey information is not always perfect. Some types of problems like invalid values, contradictory answers to related questions or non-response are quite frequent and have to be taken into account when analyzing the data. Sample size and different modes of survey data collection (e.g. CATI vs. PAPI) are some of the explaining factors for inconsistent data from different surveys. It is therefore rather the rule than the exception that different national data sources provide different data for the measurement of the same migratory phenomenon.

34. While the designation of a survey as the only data source for a certain kind of information can avoid confusion among users, it may also be argued that different data sources serve different purposes. In Switzerland, the larger sample of the structural survey (200,000 as opposed to 126,000 in the SLFS) provides data for smaller population sub-groups or smaller geographical units (e.g. NUTS 3). The SLFS, on the other hand, enables to cross-classify migration characteristics with a myriad of socio-economic variables.

IV. Conclusions

35. Analyses of 2010-2012 data provided empirical evidence of some of the issues to be confronted with and the challenges related to the use of currently available Swiss migration data.

36. A key aspect of the new Swiss census system is the linking of data from population registers with data from an annual sample survey via the social security number. Information from both sources are thus combined and supplement each other. As a result, migration-relevant information is available more frequently and on a wider range of topics.

37. Stock data (either from traditional censuses or survey questions) provide more information than only the socio-demographic characteristics of population groups relevant

to migration who are present in a country at a given moment of time. They can also inform on the previous place(s) of residence, therefore enabling to identify (some) immigrants who arrived during a specific period of time in the past and contributing – to a certain extent – to the analysis of recent immigration. However, stock data are not particularly well suited for producing comprehensive data on migration flows as figures on in- and outflows are needed on a continuous basis in order to capture all relevant movements. Administrative sources are frequently able to do so despite some of their well-known drawbacks (e.g. statistical data collection is not a priority, definitions and coverage depend on legislation and administrative rules).

38. Sample surveys may not cover the total resident population, thus excluding certain population groups relevant to international migration. Efforts could be made to collect missing information either through expanding the survey universe or – if not feasible – identifying and exploiting alternative sources.

39. Since there is no single data source which provides a comprehensive picture of migration it is good practice to use multiple sources. While in some instances these sources will lead to the same findings, in other instances the details and trends may differ. However, this does not necessarily mean that one source is 'right' and another is 'wrong', or that one source is 'better' than another one. As a consequence, it may be useful to prioritize sources according to the required information (e.g. at a certain geographical level or in combination with other variables inherent to a particular survey). On the other hand, combining data from different sources to provide a coherent picture can be methodologically challenging due to variations in concepts and definitions.

40. Further work will be done at FSO to close some of the existing gaps and to improve the quality and availability of migration data. A specially designed national migration module will be attached to the Swiss Labor Force Survey of 2014 to complement the EU LFS ad hoc module on the labor market situation of migrants and their immediate descendants and provide previously unavailable data. Finally, a recently approved new National Centre of Competence in Research 'On the Move: The Migration-Mobility Nexus' will be devoted to the academic study of migration-related topics – with special emphasis on interdisciplinary and new innovative approaches and with substantial funding by the Swiss National Science Foundation. FSO will be closely involved in some of the research projects.

United Nations ECE/CES/2014/28

Economic and Social Council

Distr.: General
29 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration statistics

Better monitoring migrant populations and international mobility: progress made, challenges and ways forward

Prepared by the Organisation for Economic Co-operation and Development

Summary

This paper provides an overview of the work undertaken by the Organisation for Economic Co-operation and Development in order to address some of these challenges and to better respond to policy needs. In particular, the most recent update of the Database on Immigrants in the member countries of the Organisation for Economic Co-operation and Development, based on the 2010-2011 round of population censuses, allows measuring the main changes in the composition and characteristics of migrant populations and provides new estimates for emigration rates by skill level for more than 180 countries of origin worldwide. Furthermore, efforts have been undertaken to provide more up-to-date flow data in order to support evidence-based policy making in the field of migration management. This paper also addresses briefly some of the grey areas and potential improvements in terms of measuring migrant populations and international mobility.

The paper is presented for discussion to the Conference of European Statisticians seminar on Migration statistics.

GE.14-20514

1420514

Please recycle

I. Introduction

1. Between the creation of the Organisation for Economic Co-operation and Development (OECD) in 1961 and the Great Recession, three periods can be distinguished concerning international migration movements in the OECD area. The first period is one of increasing temporary labour migration ("guest worker" migration). It ended in 1974 with the oil crisis, and opened on to 15 years of relative stability of migration flows. And between the fall of the Iron Curtain and the 2008 crisis, immigration flows to the OECD grew very rapidly, partly due to the development of free circulation in Europe. Overall, net migration to the OECD area in the 1990s was more than twice what it was during the previous decade. This has in turn drawn more attention to issues related to international migration, and this interest has never ceased to grow since then, as flows went global and high. But by the turn of the millennium, policy makers were lacking comparable international data to take informed measures and to build evidence-based migration policies.

II. Measuring migrant populations and their characteristics

2. OECD is contributing to filling gaps in several ways. The Database on Immigrants in OECD Countries (DIOC) is one of the projects the OECD International Migration Division has undertaken in this view, with the help of the OECD Statistics Directorate.

3. With the 2000 round of population censuses, virtually all OECD countries have incorporated in their census a question on the country of birth of persons enumerated, as well as on their nationality. With this information, it was possible to provide, for the first time, a detailed, comparable and reliable picture of immigrant populations in OECD countries, reflecting the cumulative effect of movements within and to the OECD area over the past decades. And with additional information on the educational attainment of migrants, the cumulative impact of flows of human capital could be depicted and, in particular, the conventional wisdom on the brain drain confronted with actual data.

A. The database on immigrants

4. The first edition of the Database on Immigrants in OECD Countries (DIOC) capitalised on this new source of information, supplemented if necessary by population registers and labour force surveys, to provide comprehensive and comparative information on a broad range of demographic and labour market characteristics of immigrants living in OECD countries. The database was compiled in collaboration with OECD national statistical offices.

5. DIOC includes information on demographic characteristics (age and gender), duration of stay, labour market outcomes (labour market status, occupations, sectors of activity), educational attainment and the place of birth. It provides the possibility of comparing immigrants both across OECD countries and with their respective native-born populations (in a large number of countries), thus shedding light on the differences and similarities between these groups. With parallel information on demographic and social variables on persons in origin countries who did not migrate, one can compare the situation of migrants with that of non-migrants on a broad range of characteristics, including educational attainment and labour market outcomes, thus improving our understanding of the migration-development nexus. It can also be used to address policy-related questions

related to the brain drain, the gender aspect of migration, health professionals' mobility or the role of low-skilled foreign-born workers in the service sector.

6. An update of DIOC based on 2005/06 data was produced in 2011. Given the smaller number of countries which held a mid-decade census, labour force surveys were more widely used for this edition.

7. OECD and the World Bank also joined their efforts in a project aimed at extending the coverage of the Database to non-OECD destination countries (DIOC-E or DIOC extended). This data collection made it possible to calculate emigration rates by skill level for more than 180 countries, and covered about three quarters of the migrants worldwide.

8. In the first quarter of 2014, the third edition of the Database will be released. Based on the 2010-11 round of population censuses, it allows measuring the main changes in size, composition and characteristics of migrant populations and provides new estimates for emigration rates by skill level for over 180 countries of origin worldwide. This edition also contains new information on international migration of teaching and health personnel and field of study. An updated version of this Database will be released at the end of 2014 and will include a number of non-OECD destination countries. This work is conducted in cooperation with the World Bank and the International Migration Institute at the University of Oxford.

B. Statistical challenges and limitations

9. While developing this database, a number of statistical issues have emerged.

10. First, many countries do not release information on small groups (cells that are below a certain size). The objective of this database is to provide a complete profile of emigrant populations, including for small countries. Hence considerations of this kind are important ones. In such cases, to address countries' concerns about confidentiality, the technique of random rounding has been used. In addition, imputation of some data has been necessary in the cases of countries which have not been able/ allowed to provide detailed information for all countries of origin.

11. Second, a major challenge is related to differences in the classifications used across countries. This is the case with classifications used to measure the educational attainment, where matching problems were mostly observed in the first edition of the database. These issues have been tackled by proactive work with the countries statistical offices at the earliest stages of the production.

12. Similar difficulties appeared around classifications by occupation. For example, the classifications used in the United States and in Mexico cannot be matched with the international classification, used by other OECD countries. In addition, information is not available at the same level of detail in all countries. Some countries do not provide 2-digit detail, but only 1-digit, which limits the scope of potential analyses. The database addresses this issue, by including various variables on occupation based on the different national classifications or aggregation levels.

13. An additional challenge concerns the variables provided by the different countries. For example, not all countries have been able to provide information on the years of residence or the field of study, or some countries were only able to provide aggregate information on these variables.

14. Finally, concerning the sources, the main issue is that census data is not available for all countries, in which case labour force surveys (LFS) are used instead. LFS have smaller sample sizes and do not allow detailed disaggregation especially in terms of occupation,

educational attainment and years of residence (also field of study?). This is particularly an issue in the 2005-06 update of the Database which is largely based on LFS data.

C. First results of the database on immigrants

15. The first global results drawn from the database for 2010-11 give a quantitative measure of the extent to which the economic crisis has affected migrant populations and population movements. Please note that the Database is currently being finalised so the results presented here may differ slightly from those which will be published in March. However, the global picture should remain quite similar.

16. In 2010-11, about 100 million persons aged 15 years old and over were living outside their country of birth in the OECD. This represented a 36% increase from its 2000-01 level and a 9% increase in comparison with 2005-06. The slower growth of migrant stock in the OECD area in the second half of the decade reflects the decline in migration flows due to the economic crisis affecting many developed economies from 2007-08 as well as a higher return migration. Women represent 52% of the immigrant population, up from 51% in 2000-01.

17. Europe is still the main region of origin of migrants in OECD countries in 2010-11, but despite the EU enlargement waves of the decade, its share declined from 35% to 33% over the period. Asia is the second region of origin with 27%, up 2 percentage points since 2000-01, closely followed by Latin America and the Caribbean, stable at 26%. Africa is the other region having seen its share increasing: 11% against 9% ten years earlier, which represents a 53% growth of the African-born population aged 15 and above living in OECD countries.

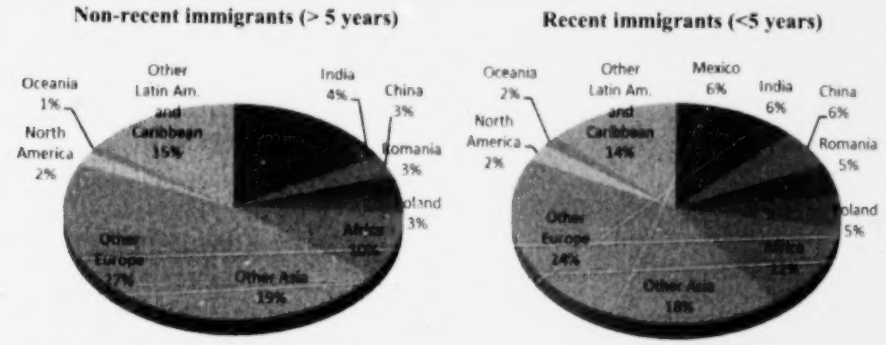
18. Mexico remains the main country of origin with over 11 million emigrants living in OECD countries in 2010-11. China takes the second place with 3.8 million emigrants, moving up two ranks compared to 2000-01 and India jumps from ninth to fourth place (3.4 million).

19. The number of highly educated immigrants in OECD countries rose sharply in the past decade, from 16 million in 2000-01 to 28 million in 2010-11, representing 29% of the total.

20. Examining the migrants' duration of stay allows to better understand how the composition and characteristics of migrants changed with the economic crisis. As can be seen in Figure 1, the recent immigrants, i.e. those arrived in the five years between 2005/06 and 2010-11, are substantially different from the non-recent ones (those who were already in the host country in 2005) in terms of distribution by country or region of origin.

Figure 1.

Distribution of immigrants by duration of residence and by place of birth, 2010-11

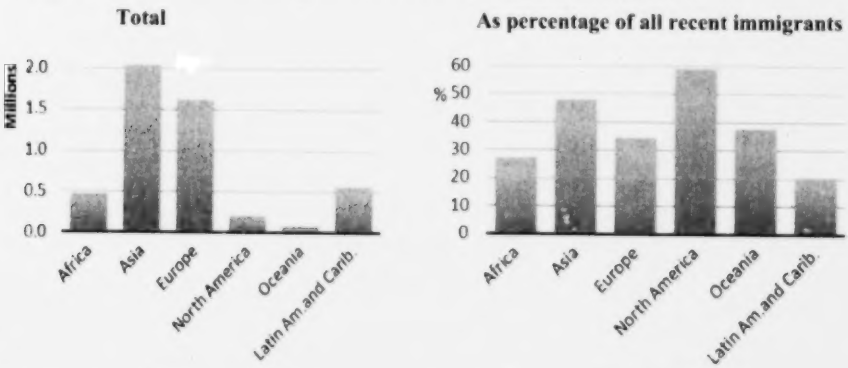


Source: OECD Database on Immigrants in OECD Countries (DIOC) 2010-11

21. If the main country of birth of recent immigrants in OECD countries is Mexico, the share of Mexican-born among recent immigrants is only half what it is among the non-recent immigrants. By contrast, the share of immigrants born in India, China, Romania or Poland is almost twice higher among the recent immigrants than among the non-recent. One in eight recent immigrants was born in Africa, against only one in ten of the non-recent immigrants. By contrast, the share of recent immigrants born in the rest of Europe (other than Romania and Poland) is three percentage points below what it is among the non-recent ones.

Figure 2.

Highly educated recent immigrants to OECD by region of birth, 2010-11



Source: OECD Database on Immigrants in OECD Countries (DIOC) 2010-11

22. In 2010-11, five million tertiary educated persons had migrated to the OECD area in the five preceding years. Two million of these recent highly skilled migrants were born in Asia, accounting for almost half of all recent migrants from Asia. 1.6 million were born in Europe, around 600 000 in Latin America and the Caribbean, and half a million in Africa.

23. Unsurprisingly, migrants have been severely affected by the economic crisis. Their global unemployment rate was 9.6% in 2000-01 and remained unchanged in 2005/06 but was 2 percentage points higher five years later. Migrants originating from Africa were hit

hardest, as their already high unemployment rate went from 15% in 2005/06 to almost 20% in 2010-11. This increase in unemployment is explained by the presence of African migrants (especially men) in OECD countries (the United States, Spain, Italy, etc.) and sectors heavily hit by the crisis (construction and manufacturing).

24. Emigration rates to the OECD area are higher than ever before, notably for Latin America and the Caribbean. Indeed, almost 6% of all persons born in this region were living in OECD countries in 2010-11. Small countries are also particularly exposed to "brain drain". The share of highly educated persons born in a country and living in the OECD area is notably high in a number of Caribbean countries. As a region, Africa is the most concerned with one in every nine tertiary educated persons born in Africa residing in an OECD country. Corresponding figures for Latin America and the Caribbean, Europe and Asia were one in 13, one in 20 and one in 30, respectively.

25. "Brain drain" is more acute for women than for men. To illustrate, in more than six out of ten countries of origin, the share of tertiary educated women who were living outside their country of birth was higher than for men. In a number of countries, mainly in Africa, the difference even reached 10%.

26. If the highest emigration rates to OECD countries are recorded for small countries and island states in the Pacific or the Caribbean, the largest increases in emigration rates since 2000-01 are observed in European countries in the Balkans or the Baltic States.

D. A unique tool

27. This Database has proven a unique tool in analysing the size and characteristics of emigrant populations by country of origin. The addition of the 2010/11 data allows for an analysis over a decade but also includes a number of improvements and developments, such as the extension of the fields covered (e.g. emigration of education professionals). This work will continue with a view to further improve the quality and the relevance of the database for analytical and policy purposes.

III. Measuring international mobility

A. Standardised statistics on international migration

28. At the end of the 90s, the growing migration flows to OECD countries and the diversification of destination countries urged for in-depth reflexions on migration flow definitions and international comparability. At this time, official data did not generally include information on the nature of the migration flows, that is, whether they are for work, family, humanitarian or other reasons. Empirical results have shown that integration outcomes vary widely by migration category although most long-term immigrants can enter the labour market once they have been admitted into the receiving country. The migration category is a matter of interest for policy since some categories are not subject to full discretion regarding admission by host country authorities. Cross-country differences in the distribution of migration by category in part explain differences in immigrants' economic and social outcomes across countries (OECD, 2012). In addition, many of the standard national sources tended to include some short-term movements along with long-term flows and distinguishing between these two is also of interest.

29. Building on the major methodological work published by the United Nations (United Nations 1998) and OECD (Lemaître 2005), OECD started to publish standardised statistics for most OECD countries (for an overview, see Fron et al, 2007). Three objectives

have guided most of the choices made with respect to the publication of these standardised long-term migration flow data. They are: a) international comparability; b) the need for statistics on long-term migration; and c) disaggregating data by category of migration. These three objectives together have determined the choice of residence permit data as the main data source used.

30. Permanent migrants are defined here as persons who have been granted the right of permanent residence upon entry or persons admitted with a permit of limited duration but more or less indefinitely renewable, plus entering persons with the right of free movement who are expected to stay (i.e. excluding short-term movements and students). Note that this includes status changes, i.e. persons who were first admitted on a what would be considered a temporary category and who afterwards changed category, as could be the case of asylum seekers who have seen their request accepted and received a permanent residence permit or international students who take up employment or marry a resident in the host country..

31. Another limitation of data based on residence permits is that a permit may be granted but not actually used, which may lead to some overestimation. On the other hand, it is possible that an immigrant may be granted more than one permit, for example a person entering for family reunification may also receive a work permit allowing him or her to work part-time and such double-counting cannot always be excluded. Also, the date when the permit was delivered may not correspond to the date of actual entry in the country.

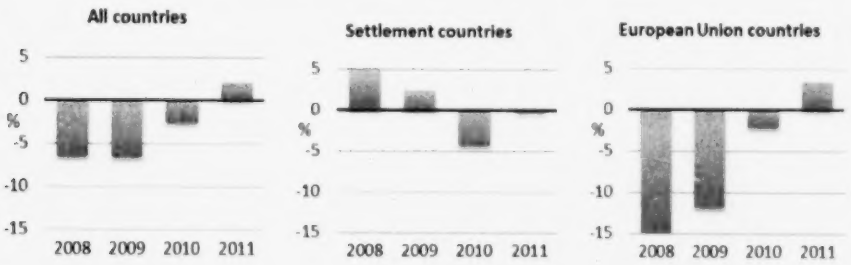
32. In the first years of existence of this dataset, flows of migration within free circulation areas were reallocated to categories such as work or family according to their labour market status, but it has proven uncertain to accurately estimate this distribution. These movements are now presented separately from regulated movements. The category of work-related migration is now restricted solely to the discretionary worker migration.

B. Major trends during the economic crisis

33. In the first years of the crisis, flows of permanent migration to the OECD have dropped significantly, by 7%, both in 2008 and in 2009 (Figure 3). European Union countries experienced a severe decline in flows, but this probably has to be weighted by the fact that flows to some European countries had increased very sharply prior to the crisis. Permanent flows to settlement countries have been much more stable over the crisis period, displaying variations comprised between +5% and -4%.

Figure 3.
Inflows of permanent immigrants into selected OECD countries, 2008-2011

Percentage change over previous year

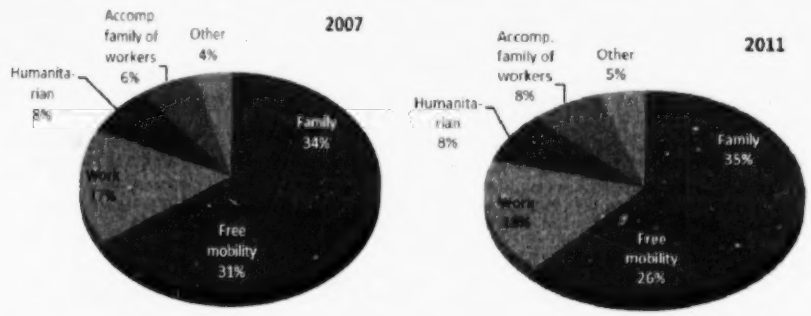


Note: All countries: 23 OECD countries covering 95% of migration flows to the OECD. European Union countries for which standardised statistics are available. Settlement countries: Australia, Canada, New Zealand and United States.

Source: OECD International Migration Outlook 2013 – Standardised statistics.

34. Overall, the composition of the permanent flows to OECD countries reveals that the most frequent reason for migration is inked with family (Figure 4). In 2007 and in 2011, more than a third of the movements concerned family migration. Migration within zones of free circulation is in fact the only category which has seen its share declining during the economic crisis.

Figure 4.
Permanent immigration in OECD countries by category of entry, standardised statistics

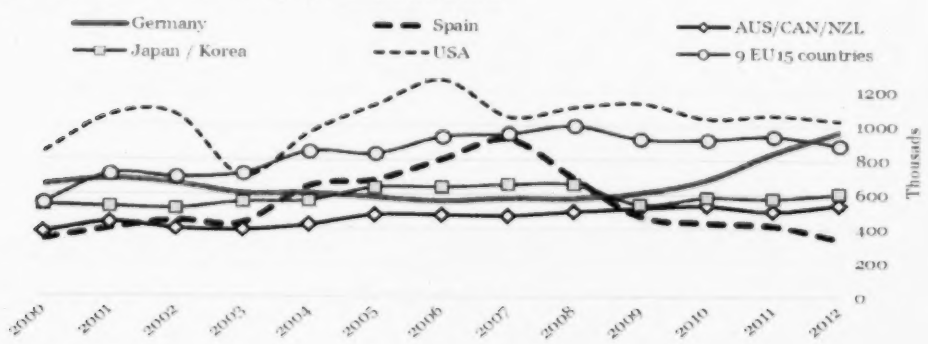


Source: OECD International Migration Outlook 2013 – Standardised statistics.

C. On-going developments

35. As can be seen in Figure 5, the magnitude of migration flows to OECD countries increased again significantly during the 2000s and particularly since 2003. Even if a declining trend has been observed since the 2007 peak, flow levels remain higher than they stood in 2000. Policy makers' interest has developed accordingly, and so has the demand for more up-to-date and detailed statistics.

Figure 5.
Inflows to OECD countries by destination, 2000-2012



36. OECD has launched a project aiming at developing further the standardised statistics on permanent flows. First objective is to improve the timeliness of the data currently published. Another objective is to extend the coverage to all OECD countries as a number of countries are still not covered – although the countries included account for 95% of the inflows to OECD. A third objective is to include new breakdowns, notably by nationality and gender.

37. For this purpose a questionnaire on data availability has been sent to the members of the OECD Expert Group of Migration in order to examine the feasibility of such developments. The results of this questionnaire allowed to carry on with the project as for example data by nationality are generally available, so are data by gender, and also, infra-annual data are available in two-thirds of the countries.

Table 1.
Inflows to OECD countries, 2011-2013, partial data

Country	2011	2012	2013	% change 2011-12	% change 2012-13	Period covered	Months
Australia	210 700	242 400	254 700	15	5	Jul-Jun	12
Canada	117 300	128 600	133 700	10	4	Jan-Jun	6
Denmark	48 200	52 900	55 500	10	5	Jan-Oct	10
Finland	23 700	21 300	21 400	-10	0	Jan-Nov	11
Germany	381 000	447 000	501 000	17	12	Jan-Jun	6
Ireland	33 700	32 100	40 200	-5	25	May-Apr	12
Netherlands	59 600	57 500	57 800	-4	1	Jan-Oct	9
New Zealand	40 700	40 400	39 000	-1	-3	Jul-Jun	12
Norway	15 900	16 600	16 600	4	0	Jan-Jun	6
Sweden	93 100	111 100	116 600	19	5	Jan-Nov	11
Switzerland	138 600	146 100	151 500	5	4	Sep-Aug	12
Total*	**1 162 500**	**1 296 000**	**1 388 000**	**11**	**7**		

Source: National sources.

38. As already mentioned, timeliness is a key objective, and in order to provide as timely as possible data and information, infra-annual data are necessary. Table 1 is an example of what OECD provided to a high-level meeting at the end of last year. This table shows a comparison of the available 2013 data on migration flows with the data for the corresponding period of the previous year. This type of information is relevant in terms of policy making, as it gives an idea of the variation. It is interesting to note that in 2012, seven of the eleven countries show a positive variation on 2011, and that for all countries but one, there is an increasing trend in 2013. However, currently, these figures available are not comparable between countries. So OECD aims at improving the timeliness while keeping the comparability of the statistics on permanent flows.

39. OECD is also about to launch a wide reflexion on the measurement of temporary workers through its OECD Expert Group on Migration. The objectives are to better capture the reality of temporary migration, in particular to deal with the status changes, and to examine which categories are growing in importance in migration policies and should be subject to heightened attention.

IV. Other topics

40. OECD is also working on several issues such as return migration, which is looked at carefully, especially in times of crisis, although it is difficult to adequately assess the magnitude of the phenomenon. The same goes for irregular migration and illegal movements, which are regularly dealt with in OECD publications and meetings.

41. A topic of high importance for a number of OECD member countries and members of the OECD Expert Group on Migration is the emigration of nationals, especially, within free circulation areas. Capturing emigration and emigrant populations' profiles can help to define policies to better mobilise the Diasporas and reduce the possible negative impact of high emigration rates, especially those of the highly-skilled.

42. Attracting and retaining skilled migrants and in particular students is a goal shared by many OECD countries. OECD has provided estimations of "stay rates" for international students. The stay rate is estimated as the ratio of the number of persons who have changed status (whether for work, family or other reasons) to the number of students who have not renewed their permits. However, this measure is not without limitations. Because the change-of-status statistics are based on permit data, they do not include citizens of the European Economic Area (EEA) for European countries, who do not need a student permit to study in another country of EEA.

V. Conclusion

43. As international migration is increasingly shaping the societies of OECD countries, obtaining better and more comparable information on migration flows and the characteristics of immigrant and expatriate populations is an increasingly important challenge. It is both a prerequisite for better informed policy making as well as for a balanced public debate in this domain.

44. Significant progress has been made in this respect over the past decade by OECD and its partners, through the Database on Immigrants in OECD countries (DIOC) and through the standardisation of migration statistics by migration category. At the same time, there remain a number of important gaps in the statistical infrastructure in terms of measuring migrant populations and international mobility that need to be addressed. These include the emigration of nationals and mobility within free circulation areas, which are among the most difficult flows to measure; the issue of return migration and the retention of skilled migrants and, finally, the evolution of irregular flows.

United Nations ECE/CES/2014/29

Economic and Social Council

Distr.: General
29 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration Statistics

International migration in Mexico and its statistics

Prepared by the National Institute of Statistical and Geography of Mexico

Summary

The paper discusses initiatives to improve the conceptual and methodological accuracy of migration statistics and the related information sources. It considers ways to increase coordination and collaboration among agencies in the move towards a coherent and comprehensive national system of statistics on international migration. The paper also makes similar proposals to census sources and data in the United States and in countries of Central America, to progressively establish better statistics on migration with a broader regional perspective.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

GE.14-20518

* 1 4 2 0 5 1 8 *

Please recycle

I. Introduction

1. International migration in the world has increased substantially in recent years. The International Organization for Migration estimates that between 1965 and 2000, migration increased from 75 million to 175 million people. Today, the number of international migrants is around 232 million, equivalent to 3.2 percent of the world population (OIM, 2013:64). Migratory displacement now takes place in parallel with the accelerated process of global integration, articulating economies and people.

2. International migration is an important process in modern life and should be integrated into development policies. It has become essential to have reliable statistics about migration. However, measuring it continues to be a challenge nearly everywhere in the world. Conceptually and methodologically, it is a complex phenomenon, given its various forms, flows and cycles.

3. Mexico's international migration stands out because of its large dimensions and varied characteristics. Around 12.4 million people born in Mexico live in another country; this means that Mexico has the second biggest flow of international migrants in the world. In addition, this migration has the characteristic of concentrating itself in one destination: the United States of America. Mexicans make up around 6.8 percent of all international migrants worldwide[1]. One in ten Mexicans lives outside of their country.

4. International migration involving Mexico, in addition to a) the flow of Mexicans to the United States of America, includes: b) the immigration of foreigners to Mexico; c) the transit of foreigners-mostly Central Americans-through Mexican territory to the United States, which turns Mexico into a migrant corridor; d) the voluntary and involuntary return of Mexican migrants to Mexico; e) the voluntary and involuntary return of Central American migrants to their countries of origin, through and from Mexico; and f) the cross-border labor flows between Guatemala and Mexico and with the United States.

5. Each one of the afore-mentioned flows has its own complexity, taking into account at least three variables: its temporary cycles; its evolution through various junctures (examples include the labor market and U.S. and Mexican migration policies); and its regular or irregular forms (documented or undocumented), with the latter being the most difficult to measure and respond to.

6. For Mexico, international migration has profound relevance. Its greatest impact has been to change our definition of nation. The nation, once only seen as encompassing the country's territory, now must adapt to its cross-border reality. International migration also has transformed the life of Mexico's regions, its economy, and its social, cultural, and political dynamics. The public policies of the state and development strategies continue to adapt to the situation. Strong challenges persist involving the protection of the human rights of Mexicans abroad, as well as protecting the rights of citizens of other countries traveling through Mexico. No less important are border dynamics advancing toward a new regionalization of development both with the United States and Guatemala, with their own respective challenges. In sum, international migration has become a powerful force of social change in practically all national areas. This is not to mention that in the United States there also are similar processes of large-scale change resulting from this migratory intensity.

[1] United Nations, Department of Economic and Social Affairs (2012). Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations database, POP/DB/MIG/Stock/Rev.2012)

7. Statistics about migration in Mexico have made significant advances and are able to measure the main components of the process; however, they still need greater development. A gap continues to exist between the national importance of migration on the one hand, and on the other the statistical and official record-keeping systems that still are unable to document the magnitude of the migratory processes. Also pending is the regional collaboration that needs to exist between the countries of North America and Central America, given that international migration links the entire region's economies and populations.

II. The international migration of Mexicans

8. The demand for unskilled labor in the United States has been the determining factor for migration from Mexico. Between 1900 and 1920, this movement was driven in part by the demand for workers brought about by World War I; this stage has been labeled the "Era of Enganche," with "enganche" defined as hook or indentured (Durand and Massey, 2003). A similar situation occurred 20 years later, with World War II. In 1942, the Bracero Program began. It was a labor agreement between the governments of Mexico and the United States that allowed the temporary hiring of Mexican males; those hired came from, and worked in, rural areas. The program expanded as a result of the postwar economic boom, bringing about a circular movement of nearly 5 million workers between 1942 and 1964, when the program ended.

9. Thus was established a large-scale, informal labor market sustained by large economic, social, and family networks. This migration has persisted in a context of strong economic asymmetry; for decades, the United States has been the main world power, while Mexico has continued along the lines of a developing country. One economic indicator is the wage differential between both countries. In 2012, the direct pay average in manufacturing wage was $27.15 per hour in the United States; in Mexico, the equivalent was just $4.45.[2]

10. Today, Mexican migration to the United States needs to be understood in the context of the economic integration of both countries, following the 1994 implementation of the North American Free Trade Agreement (NAFTA), which also included Canada. Although the agreement did not consider a unified labor market—that continues to be off the table—the fact is that trade has created strong interdependencies: Mexico is the third--most-important commercial partner of the United States (after Canada and China), while the United States is Mexico's main source of imports and its main export market.[3] International migration and its powerful social consequences in both countries, along with the two countries' economies, have created solid bonds.

11. International migration has traditionally not been a central part of the design and instrumentation of Mexico's population and housing census; thus, its measurement from this source depends on indirect methods. This situation has improved in the last two censuses (2000 and 2010); nevertheless, migration still could have a much greater presence in the surveys. The censuses and another set of national surveys are conducted by the National Institute of Statistics and Geography, INEGI (www.inegi.org.mx).

---

[2] U.S. Bureau of Labor Statistics (BLS), http://www.bls.gov/fls/ichcc.htm#table2.2

[3] In 2012, 16.1 percent of foreign trade was with Canada, 14 percent with China, and 12.9 percent with Mexico. [United States Census Bureau, Top Trading Partners, http://www.census.gov/foreign-trade/statistics/highlights/top/top1212yr.html]

12. For previous decades, indirect methods were used to calculate the number of emigrants, as they are the combination of the residual method and the book-keeping equation. The difference between the total population is calculated, at two points in time, and the increase is compared with natural population growth. This residual procedure reveals net international migration. The following estimates were made: between 260,000 and 290,000 from 1960-1970; between 1.2 million and 1.55 million from 1970-1980; between 2.1 million and 2.6 million from 1980-1990; and 1.387 million in the five-year period between 1990 and 1995 (Corona and Tuirán, 2000).

13. Beginning with the 2000 Census, a special form was incorporated to compile information about international emigration (involving 2.9 million households). This information showed the territorial dispersion of emigration. Until the 1970s, emigration had its origins in a small number of Mexican states (seven of thirty-one) known as "traditional sending regions." In the five years between 1995 and 2000, 47 percent of Mexican emigrants originated from these regions; this percentage fell to 40 percent during the 2005-2010 period (CONAPO, 2012). This shows that "non-traditional" sending regions now are the new sources of emigration.

14. Using census information, the National Population Council, CONAPO (www.conapo.gob.mx), developed an index of the intensity of migration for each state and for Mexico City (Federal District). This appears in Graph 1, which describes the main regions from which Mexicans migrate internationally.

Graph 1

States in Mexico according to their intensity of international migration, 2010

Source: Self-generated with data from the National Population Council (CONAPO).

Note: An entity showing a value of zero would have no intensity of migration at all; one showing a value of 100 would mean that each of four indicators was 100 percent. None of the federal entities met these conditions.

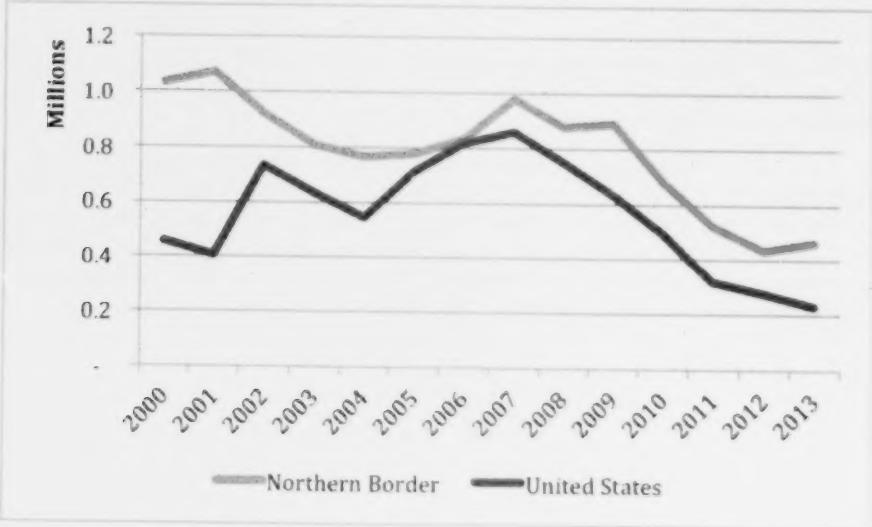
A. Surveys

15. The most specialized is the Survey on Migration at the Northern Border of Mexico (EMIF Norte), devoted to the direct study of migrant flows where they occur, and taking into account the direction of the flows.[4] This survey is unique in Mexico (and possibly in the world) because its methodology is designed to catch people in motion, including formal and informal movement. The result is a migration observatory, with quarterly statistics that include socio-demographic characteristics, migration and work trajectories, reasons for emigration, conditions of migrants' movement, the origins and destinations of the migrants, and the risks of the border crossing, among other important aspects.

16. The results show that between 2000 and 2007, the flow of migrants to the United States showed a continuing increase, only to later begin a pronounced drop, settling at a relatively stable number in 2013. Whereas in 2000 more than 450,000 migration events to the United States (with or without crossing documents) were measured, in 2007 this number grew to a little more than 850,000. In 2013, the number fell to an estimated 270,000. (Graph 2 also shows population movement to northern Mexico border-city destinations).

Graph 2

Migrant flows from the south by destination, 2000-2013



Source: EMIF Norte, flow of migrants coming from the south 2000---2013. The data for 2013 only goes through the third quarter.

[4] It began in 1993 as a joint effort between El Colegio de la Frontera Norte (College of the Northern Border, or COLEF), the Labor and Social Welfare Ministry (STPS), and the National Population Council (CONAPO) in order to provide reliable statistical information about migration between Mexico and the United States. After 2003, the National Institute of Migration (INM) and the Foreign Ministry (SRE) joined the project. Its main databases and tabulated results can be consulted at http://www.colef.mx/emif/.

17. The Migration Survey in the Northern Border (EMIF Norte) describes the dimensions of the labor market between Mexico and the United States, as well as its cycles. It is a large-scale market, built over decades, but not recognized in bilateral relations. This means that there is a large amount of informality; as a result, there are vulnerabilities of all types for the migrant population. The U.S. recession lowered demand for unskilled labor, reducing the movement of workers coming from Mexico, particularly since 2009 began. Additionally, this coincides with a period where migration policies have hardened and where barriers and walls along the border have been reinforced, additional factors inhibiting circular labor migration. More aggressive deportation policies affecting Mexicans and people from other countries also have been implemented.

18. The EMIF Norte also registers return flows, including migrants coming from the United States (voluntary returns) and those repatriated by U.S. authorities. Large-scale deportations have been part of the structure of migration. In 2000, the highest point was reached, with more than 800,000 migration events; between 2010 and 2012 the average was above 370,000. Although the volume was reduced, the situation has been more critical in recent years due to the profile of those repatriated: increasingly they are residents of the United States, with families and work and social networks formed over years.

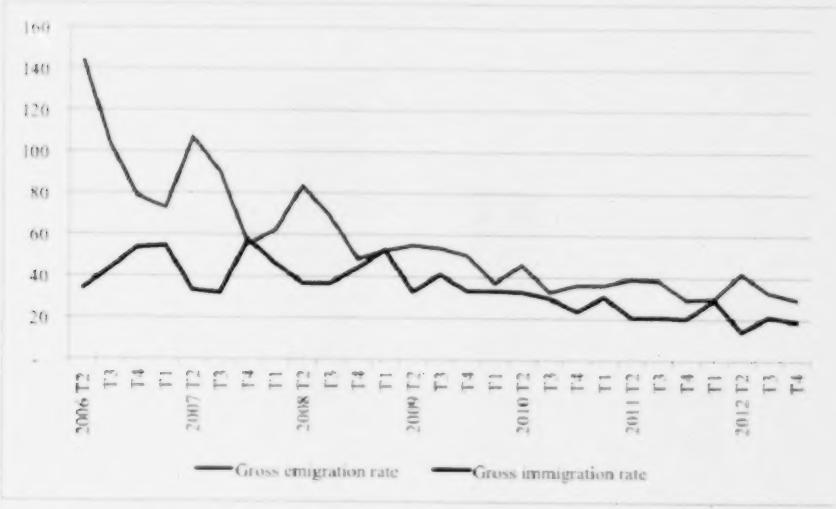
19. Until 2000, more than 80 percent of deportees were people with less than a day's stay in the United States, essentially detained upon their entry; in 2012, this number was just 15 percent. This shows that most of the recent repatriations are people who had residence there; their deportation has generated numerous cases of family separation, including fathers and mothers leaving their children in the United States, irrespective of their ages. The new policy of deportations is one of the most serious issues confronting Mexicans' international migration.

20. An additional resource to measure international migration is the National Survey of Occupation and Employment (ENOE) conducted quarterly by INEGI, conducted in households.[5] By comparing the composition of households, it is able to track the international movements of Mexicans. Results indicate that from 2006 to 2012, migration has fallen significantly: Whereas at the beginning of 2006 there were 140 migrants per 10,000 residents, in the last quarter of 2010 the figure was just a quarter of that.

[5] This information can be seen at:
http://www.inegi.org.mx/est/contenidos/Proyectos/encuestas/hogares/regulares/enoe/

Graph 3

Gross rates of international migration per quarter, 2006 to 2012 (per 10,000 inhabitants)



Source: INEGI. Estimates based on the National Survey of Occupation and Employment (ENOE), 2006 to 2012.

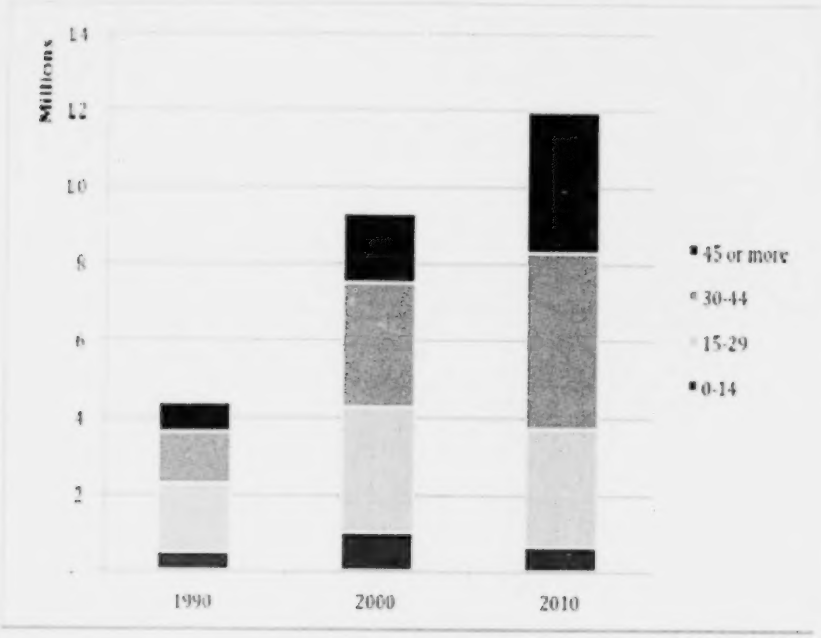
21. The ENOE also produces a migratory net balance that shows the same trajectory: The strong emigration that took place from Mexico during the 1990s until 2007 has been reduced. The net balance went from minus 110 migrants (per 10,000 residents) in 2006 to minus 10 in 2012, even showing values very close to zero during some three-month periods. The drop in the migratory net balance in recent years is the most important indicator of the current migration situation: The intense movement of Mexicans toward the United States has stopped, at least temporarily, in comparison with previous years.

B. Estimates from U.S. sources

22. The most extensive resources to measure migration from Mexico to the United States are the Current Population Survey (CPS) and the American Community Survey (ACS), of the U.S. Census Bureau. Despite the precision of both surveys, the measurement of Mexican migrants faces limitations because large numbers of them are undocumented, bringing about a large risk of underestimation (Massey, 2004:1077, Warren and Passel, 1987). In a preliminary document, INEGI estimated that the figure is underestimated by 12.5 percent. Despite this bias, the information in both surveys is excellent. Graph 4 presents ACS figures for 1990, 2000, and 2010. It shows that the Mexican population almost tripled during the period and that it is mainly composed of individuals in the economically active age cohort.

Graph 4

Mexican emigrants in the United States, according to age groups, 1990, 2000 and 2010



Source: Estimates based on the American Community Survey, 1990, 2000, and 2010

23. The Pew Hispanic Center, mainly using the referenced surveys, reported the following evolution of Mexican migration: around 1 million in 1970, 9.8 million in 2000, and 12.5 million in 2007, when the upper limit was reached. In 2012, the number was 12.4 million.

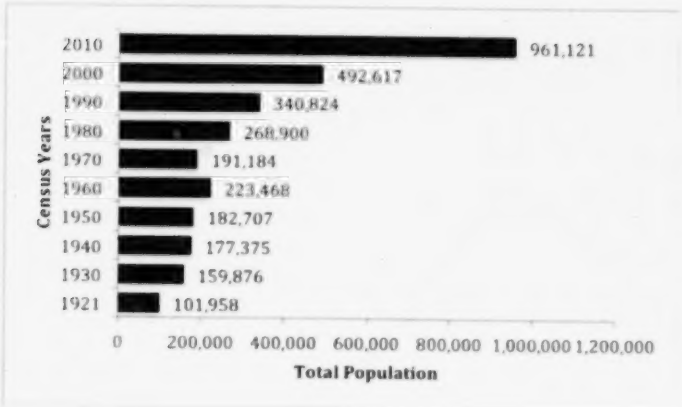
24. The sources reviewed up to this point all define Mexican migration as a large-scale process, concentrated in a relatively short period, and directed toward a single country. They agree that we are now in a period of relative stability, which seems odd following two decades of intense interchange. Net migration is tending toward zero, making it possible that the bilateral relationship could come up with a regulatory and civilized framework to define new parameters for the flow of labor.

III. International immigration in Mexico

25. The movement of foreigners into Mexico has been insignificant historically. Although its rate has grown in the last decade, particularly in the case of foreigners coming from the United States, it represents just 1 percent of the total population. It is not a relevant dynamic from a quantitative point of view, although it has some particular concentrations, such as in the northern border region. It is a migration increasingly influenced by the population of Mexican origin in the United States, and their descendants. Based on Mexican census data, it was calculated as 341,000 migrants in 1990, whereas in 2010 this number rose to 961,000, 77 percent of whom were of U.S. origin. This development and ten-year growth rates can be seen in Graph 5.

Graph 5

Foreign-born population living in Mexico, according to population censuses 1921-2010, and growth rates between censuses



C. Surveys

26. The National Survey of Demographic Dynamics (ENADID), from INEGI, includes a question about place of residence five years prior, but does not ask about birthplace. The survey questioned 575,000 international migrants returning to Mexico (ENADID 2006). This same survey, corresponding to 2009 (ENADID 2009), included birthplace, in addition to three pertinent questions: place of residence a year before; the date of return to Mexico; and place of residence five years before. ENADID 2009 thus asked 336,000 migrants coming from the United States where they had been living one year before, and 460,000 about where they had been living five years before.

IV. Transit migration through Mexico (the South-North migrant corridor)

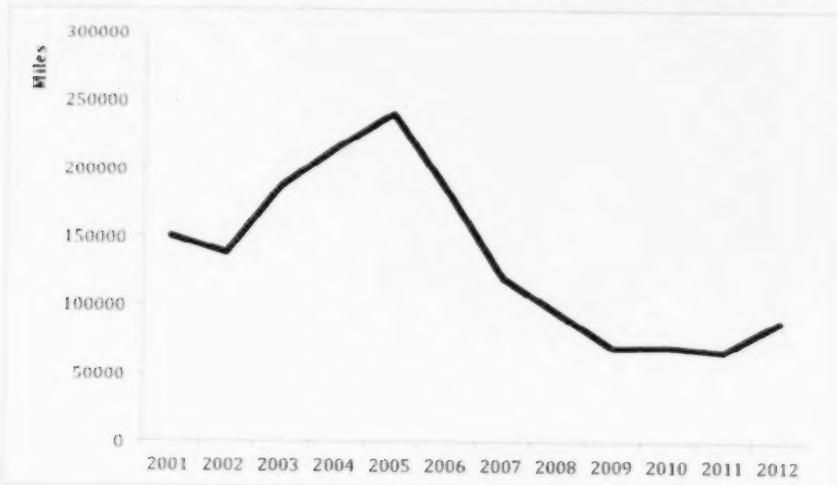
27. To estimate transit migration through Mexico is a complex task because it principally involves informal movements of people coming from Guatemala, El Salvador, and Honduras. Attempts have been made to calculate this migration using National Institute of Migration (INM) records and data from the Survey on Migration at Mexico's Southern Border (EMIF Sur).[6] However, there are large gaps in the statistics. This is the flow whose level is least known, even though good information about its composition has been gathered as a result of the involuntary return of migrants. Paradoxically, this is the most-aggrieved population in terms of human rights and is a continuing object of exploitation by organized crime.

6 The EMIF Sur has the same conceptual and methodological parameters as the EMIF Norte and is devoted to measuring labor flows at the Mexico-Guatemala border and the voluntary and involuntary return of citizens to El Salvador, Honduras, and Guatemala. It came about in 2004 as an initiative of the Colegio de la Frontera Norte, the National Institute of Migration, the Labor and Social Welfare Ministry, the Foreign Ministry and the National Population Council (http://www.colef.mx/emif/).

28. Administrative records involving in-transit migrants detained and housed in INM holding centers have been used to make an indirect count of the migrant flow. The data might reflect more the kind of job migration authorities are doing than the actual numbers of migrants in transit (Rodriguez ET to. 2011:2).

Graph 6.

Migrants housed in INM migrant holding centers, 2000-2012 (number of events)

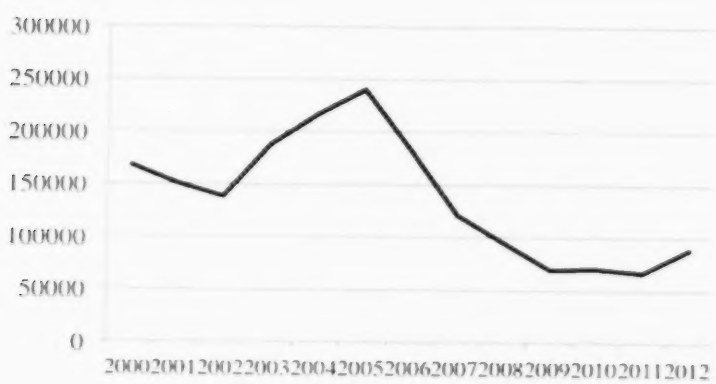


Source: Based on data from the Center for Migratory Studies (CEM---INM). Migration Policy Unit, SEGOB

29. On the other hand, the Center for Migration Studies (CMS) proposed a model—to date unique—to estimate this flow. It uses the administrative records of Mexico and the United States, as well as ACS and CPS surveys. The result is presented in Graph 7; its preparation resorted to various hypothetical scenarios.

Graph 7

Estimates of Central American Migrants (MCA) irregularly transiting through Mexico, 1995-2010

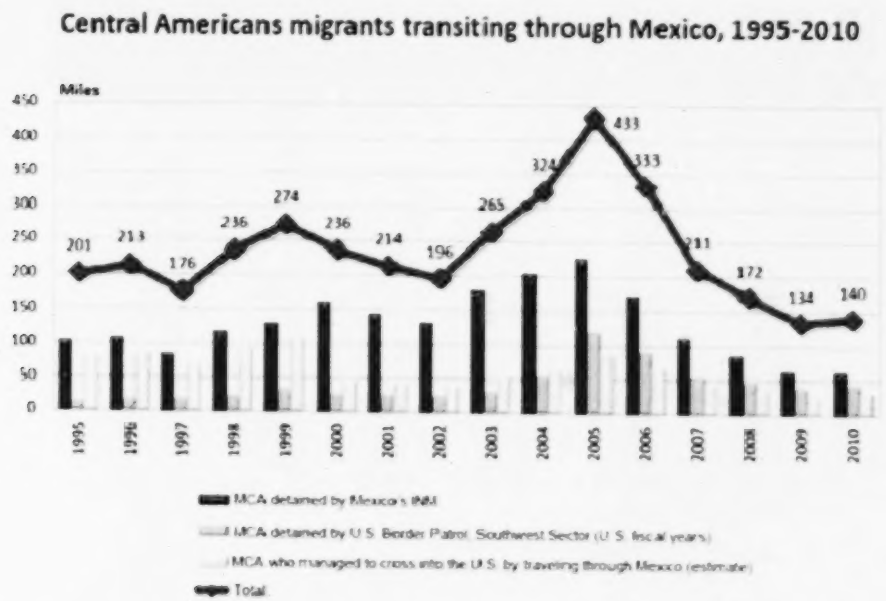


Source: Center Department of Homeland Security (Statistical Yearbook of the Immigration and Naturalization Service) and from the U.S. Border Patrol (in Rodriguez et al., 2011).

30. An additional source on the transit migration of Central Americans is EMIF Sur. The following graph only shows repatriations carried out by Mexican authorities of citizens of Guatemala, El Salvador, and Honduras, although the survey also measures deportations carried out by the United States.

Graph 8

Central American migrants headed for the U.S., returned by Mexican migration authorities, 2004-2011



Source: Based on the Survey on Migration at Mexico's Southern Border, 2004-2012

V. The cross-border workers

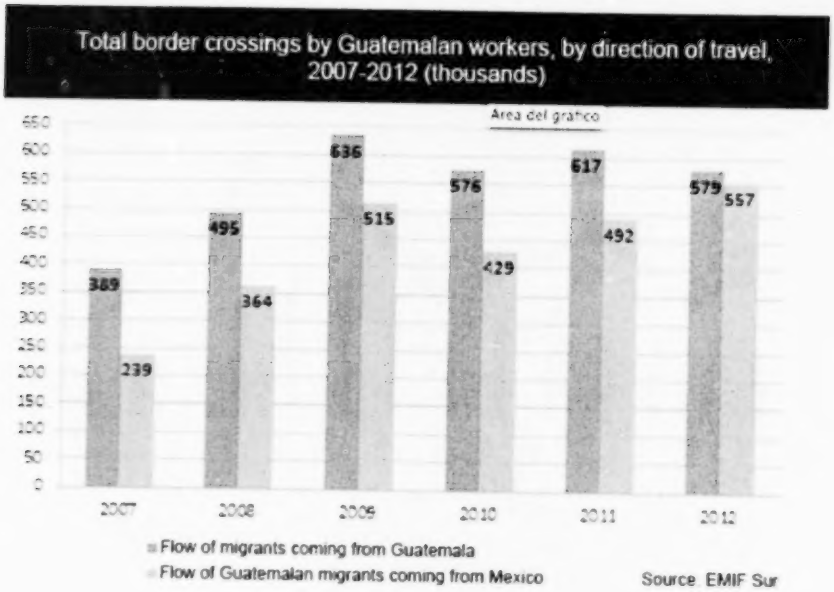
31. This migration flow is special because it passes between societies and border areas and forms part of bi-national labor, economic, and commercial networks. The people who make these daily movements have been called cross-border commuters because they must cross a border to get to their home or work (Joy, 2002). This movement takes place at both the northern and southern borders of Mexico, although in very different conditions and with very different characteristics.

32. The number of cross-border commuters in northern Mexico has been growing. In 1996, there were 28,656 of them in Tijuana, or 7.5 percent of the city's economically active population; in Ciudad Juárez there were 17,279, or 4.6% of the economically active population (Alegría, 2000). Two years later, Tijuana had 35,943 cross-border commuters, whereas in Ciudad Juárez the number fell to 15,164 (Alegría, 2002). The only source that documented the movement of these cross-border workers was the National Survey of Urban Employment (ENEU), which was conducted between 1987 and 2004. This survey was replaced by the ENOE in 2005, but it no longer included in its design the cross-border labor market; as a result, the measurement of migratory movements was further weakened.

33. At the southern border, the cross-border workers were mainly coming from Guatemala to work in Mexico's southern region; a large majority of them have papers allowing them to cross, and they increasingly have work permits. The main source of information is the EMIF Sur which observes the cross-border workers' characteristics, social profiles, details about their work, the economic sectors they work in, how long they stay in Mexico, and their various destinations, among other variables.

Graph 8

Central American migrants headed for the U.S., returned by Mexican migration authorities, 2004-2011



34. These flows have acquired a very important dimension, forming part of a bi-national region which requires greater attention to be paid to its development. At Mexico's northern border, around 850,000 migration events were recorded in 2007, the year that saw the greatest Mexican movement to the United States. At the southern border, the recorded movement of Guatemalans is around 600,000 events annually. It is worth mentioning that Guatemalan workers in the Mexican south face inadequate working conditions, including pay, social protection, and other matters.

35. An additional source of information about these links is INM record-keeping about border visitors. From 2001 to 2007 the INM distributed more than 31,000 local visitor migration forms (FMVL). Since 2009, migration policy has recognized the dimensions of these movements and record-keeping was improved; this explains the important growth of these data in recent years. There were 133,942 local border visitor permits issued in 2009, and 66,936 in 2012. Despite the drop, the records still show the intensity of this cross-border exchange and its growing formality.

VI. Conclusions

36. International migration in Mexico, as has been seen, has a number of components, some of large scale, such as the movement of Mexican workers to the United States. The determining factor—just as in other parts of the world—has been of an economic nature. It

is a huge social dynamic that has a positive impact on both countries' development, but one that falls outside of the institutional framework. For Mexico, for example, the clearest benefit is remittances, which total more than $20 billion annually; for the United States, it is to have access to workers who generate wealth in sectors such as agriculture, construction, and services.

37. To a great extent, migration continues to be informal, imposing costs such as having

to make high-risk border crossings. Between 1998 and 2012, 5,595 people died during

attempts to provide more income to their families (NFAP, 2013). With the hardening of the border and its barricade of walls and security systems, deaths have increased even though the flow has diminished (there were 477 deaths in 2012).

38. To a great extent, the dynamic of migration flows is the result of the labor market and its own dynamic. Migration policies and security systems that try to regulate (or prevent) migration are "external factors" that no doubt alter the flows, but only in a secondary manner. Or, they generate other flows that can exact severe human costs, such as repatriations, especially when they cause the separation of families.

39. On the other hand, Mexico has become a corridor for the migration of Central Americans who also go to the United States. They also face high-risk conditions in their transit through Mexico, in addition to the risks of crossing the border into their country of destination. Mexican regulations on migration have not completely adjusted to this reality; as a result, irregularity remains a common characteristic of these movements.

40. At this juncture, we find ourselves before an international migration unregistered in the formal frameworks of the states involved: While it is marginally included, it is predominantly excluded. Ideally, the United States, Mexico, Canada and the countries of Central America would come to a regional agreement, as much to regulate and to formalize this labor market as to promote policies of development to reduce asymmetries.

41. These national inequalities must be reduced in order to develop lasting alternatives to international migration and its current conditions.

42. The measurement of migratory dynamics is a complex task, in some aspects impossible to measure, given the large informal component of migrants' movements. Records concerning the origins, travel courses, and destinations of the migrants' movements are insufficient.

43. Although general measurements can be made, large gaps remain, particularly when policy requires identifying social needs at the level of specific groups and individuals. In view of this situation—and recognizing that the nation-states will not be making important changes in the short term— ideally their statistical and administrative-records systems should agree to the creation of a joint "observatory" of the widest possible scope to oversee an information-gathering system to measure migratory flows and social conditions in the most precise way possible.

VII. References

Alegria, Tito. "Demand and Supply of Mexican Cross-Border Workers." Journal of Borderlands Studies 17.1 (2002).

CONAPO. Indices de intensidad migratoria México-Estados Unidos 2010. Mexico City: Consejo Nacional de Población y Vivienda (2012). http://www.conapo.gob.mx/swb/CONAPO/Indices_de_intensidad_migratoria_Mexico-Estados_Unidos_2010.

Corona, Rodolfo, and Rodolfo Tuirán. "Fuentes mexicanas para el estudio de la migración México-Estados Unidos". Migración México-Estados Unidos: Continuidad y Cambios. Tuirán (coordinator). Mexico City: Consejo Nacional de Población (2000): 35-44.

Durand, Jorge and Douglas Massey. "Los enfoques teóricos: una síntesis". Clandestinos: Migración México-Estados Unidos en los albores del siglo XXI. Mexico City: Miguel Ángel Porrua/UAZ (2003).

Massey, Douglas, and Chiara Capoferro."Measuring Undocumented Migration." International Migration Review 38.3 (2004).

National Foundation for American Policy. How Many More Deaths? The Moral Case for a Temporary Worker Program. Washington: NFAP (2013).

International Organization for Migration. Informe sobre las migraciones en el mundo 2013. El bienestar de los migrantes y el desarrollo. Geneva, Switzerland: International Organization for Migration (2013).

Rodríguez Chávez, Ernesto, Salvador Berúmen, and Luis Felipe Ramos. "Migración centroamericana de tránsito irregular por México. Estimaciones y características generales." Apuntes Sobre Migración [bulletin of the Centro de Estudios Migratorios (Center for Migration Studies) of the INM] (2011): 1-8.

Warren, Robert and Jeffrey S. Passel. "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census." Demography 24.3 (Aug. 1987): 375-393.

United Nations ECE CES/2014/30

Economic and Social Council

Distr.: General
28 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration statistics

Measuring cross border mobility between Estonia and Finland with Mobile Positioning Datasets

Prepared by Statistics Estonia and Statistics Finland

Summary

The paper provides a statistical overview of people's cross border commuting between Estonia and Finland based on passive mobile positioning data. Studying movement of mobile phones between mobile networks enables making generalisations on the flow of people between the countries. The privacy of individuals is strictly protected according to the European Union data protection regulations. The paper also explores existing problems of measuring labour commuting between the two countries, and discusses the methodological and legal aspects of using mobile positioning data for the measurement of mobility and commuting.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

GE.14-20442

1420442

Please recycle

I. Introduction

1. The rapid increase in the spatial mobility of people is a phenomenon of the 21st century (Sheller & Urry 2006). People have become more mobile because borders have opened, tourism has become a part of lifestyles and the labor market has changed. Transnational lifestyle, living and working in different countries, has spread quickly all over the world (Massey et al 1994). The mobility of society is also significantly influenced by the development of information and communication technology (Kellerman 2013). Studying this mobile society is more complex than ever before. In a world with open borders, it is more difficult to obtain border crossing information, and transnational jobs are often connected to illegal tax schemes (Tyldum & Brunovskis, 2005). An important problem hindering the gathering of mobility statistics is, however, the fact that people commuting between two countries and places of residence often do not know themselves where they would like to live and what their perspective for the future is (Zaiceva & Zimmermann, 2008; Levitt & Jaworsky, 2007). Due to this, the lives and work schedules of cross-border commuters are unclear and they are not able to provide sufficient information about their statuses during the course of censuses and interviews.

2. One example of transnational lifestyle can be found at the Estonian-Finnish border. The mobility between these two neighbouring countries is heavy because of tourism, businesses and labour commuting. Because of open borders and the illegal labour market it is hard to control or to gather statistics about cross border mobility.

3. Labour mobility is largely caused by differences in wages between countries. These differences may be formed as a result of discrepancies between the supply and demand of labour (Russell 1995). As wages in Finland are significantly higher than in Estonia, labour mobility between Estonia and Finland is remarkable and requires more research. There is need for new data sources for compiling statistics. These are needed for developing policy measures on cross border commuting. Traditional methods are not effective in describing the diversity of cross border mobility practices (Ahas 2012).

4. Estonia and Finland are using anonymous passive mobile positioning data in order to study spatio-temporal mobility between the two countries (Ahas et al. 2008; Ratti et al. 2010; Young et al 2012). The study of movement of mobile phones between mobile networks enables estimating the flow of people between countries and regions. The privacy of individuals is strictly protected according to European Union (EU) data protection regulations. The statistical offices follow the requirements specified in EU directives on processing personal data (Directive 95/46/EC) and on the protection of privacy in the electronic communications sector (Directive 2002/58/EC). Furthermore, the Estonian Data Protection Inspectorate has evaluated the related data collection practices.

II. Data and available results

A. Passive mobile positioning data

5. Passive mobile positioning data are generated as a part of Mobile Network Operator's (MNO) operations and are automatically recorded in log files as secondary data of phone use. The most commonly available are Call Detail Records (CDR) and Data Detail Records (DDR). Different operating systems (Ericsson, Nokia-Siemens, Motorola etc.) have different technical solutions and terminology for such data outlets, but most operating systems have a similar structure (Dahlman et al, 2011; Smoreda et al. 2013).

6. The advantages of CDR data over traditional methods used to measure border crossing statistics, such as censuses, surveys and counters, are the better spatial and temporal resolution of data, a longer observation period, and automatic digital data collection. In EU, MNOs have to comply with the EU Directive 2006/24/EC for data retention.

7. In this study, we used CDR data from the memory files of the biggest Estonian mobile operator, EMT, who holds a market share of 40–45% in Estonia. The statistical offices make use of data about the use of Finnish telephones (roaming) in the EMT network for incoming visits and the use of EMT phones in Finland for outgoing visits. The management of data and the segmentation of visitors according to space-time flows is handled by Positium LBS, who used the special software Positium Data Mediator (Positium 2013). For Estonian phones abroad, data on age groups and gender from the operator's database is used and the home anchor point in Estonia is calculated using an anchor point model (Ahas et al. 2010). Because of data protection and privacy regulations, it is not possible to position phones as precisely in Finland.

8. CDR location information is initially stored as Cell Global Identity (CGI) which has a unique identity and location for each antenna. Spatially, CGI is the area of the antenna mobile radio coverage, which in reality is not physically fixed. Typically, the Global System for Mobile Communications (GSM) network's maximum theoretical signal area of a cell is 35 km, but cell antennas located along coastal areas are often further amplified to provide the best possible radio coverage across the sea. All these characteristics of passive positioning data need to be considered when using such data for migration statistics.

B. Database used in empirical study

9. Databases of call activities in the EMT network were analysed for years 2009, 2010 and 2011. This included call activities made by Estonian phones in Finland, and call activities made by Finnish phones in Estonia. Call activity is any active use (incoming or outgoing; voice, text, internet, services) of a mobile phone. The database used for the study consists of call activities (time, location) of random identifies (IDs) (Table 1). The country of origin or nationality of visitors is determined on the basis of the registration country of the mobile phone. Thus, a phone registered in Estonia may be used by a person of any nationality. Still, the registration of a mobile phone shows the place where the person is spending most of their time or has strong connections with. Random IDs are not identifiable, but certain phone numbers are given the same ID during a trip to another country by the operator. Because of privacy protection requirements we worked with an aggregated dataset, thus it was not possible to study individual entries or ID-s.

Table 1.
Data recorded on the call activity used in the study.

Phone user	ID	Time	Location
Finnish (roaming) phone in Estonia	Country; Random ID	ss:mm:hh dd:mm.yy	Network cells
Estonian phones (roaming) in Finland	Random ID	ss:mm:hh dd:mm.yy	Country

C. Visitor segments

10. Based on this CDR data various statistics are generated from roaming logs to mobility description: number of visitors, number of visits, number of days, and number of

nights. In the following section we describe the following statistics (Silm & Ahas, 2012; Järv et al 2014).

(a) Visitor – unique person (mobile phone user, ID) who is travelling to another country and has engaged in call activities while abroad;

(b) Visit – unique visit to another country by a person. One visit is normally composed of two trips: one into destination and second back home from the destination. One person can make multiple visits. Visits, which are made directly from Estonia to Finland in the case of Estonians, and directly from Finland to Estonia in the case of Finns, without transit or stopping in a third country are included in the estimation;

(c) Trip – unique one way trip to another country by a person, for example trip from Estonia to Finland; or trip from Finland to Estonia;

(d) Number of days – duration of each visit in days;

(e) Number of nights – duration of each visit in nights. The formula for calculating the number of nights in one visit is: nights = days – 1.

11. Different segments of visits/visitors based on visit duration, number of visits per year and number of days spent in another country per year are identified.

1. Visits

12. Visits are divided on the basis of the length of stay:

- Transit visits – visits to another country (Finland/Estonia) for a short period (may be <3...<12 hours) and leaving to a third country on the same day.
- Visits to the destination divided on the basis of number of visited days: one day visits, 2–4 day visits, 5 or more day visits.

2. Visitors

13. Based on the total number of days spent in the other country, visitors are divided into:

- 1 day visitors
- 2–30 day visitors
- 31+ day visitors
- 183+ day visitors. According to the common definition (WTO), a visitor who is staying in other country for 183 days or more is considered a foreign labourer.

14. Visitors are further divided on the basis of the number of visits to another country into:

- one time visitors
- 2–4 time visitors
- 5 or more time visitors.

D. Results

15. The advantage of passive mobile positioning is that it creates the opportunity to analyse foreign visits of a large part of the population over a long period of time. It allows exploring the distribution of trips by the frequency or duration of visits. The duration of trips can, in turn, be calculated in terms of one trip or year or study period. The

segmentation process can also include information on the timing and geography of trips. The main results of these distributions show that there are more persons making short trips from Finland to Estonia. There are fewer people travelling from Estonia to Finland, but the total duration of their stay in Finland is much longer.

E. Duration of trips

16. Statistics on the duration of trips in the two directions (from Finland to Estonia and from Estonia to Finland) differ greatly. Only the volume of transit trips and short-term (1–9 days) trips are comparable and demonstrate the importance of short-term visits (probably for tourism and business) in travel between the two countries. The distribution of trips lasting over 10 days shows the share of commuting for employment purposes, which is what interests policy makers: 13.4% (180,487 trips) of trips from Estonia to Finland last 10+ days. Most probably this includes persons who work in Finland, but definitely Finns who have taken up residence in Estonia as well. The share of trips from Finland to Estonia which last 10 days or more is only 2.5% (57,564 trips). The number of Estonians who spend 183 days or more (defined by WTO as foreign workers) in Finland is 13,652 (4% of total visitors). In total, 1,155 Finns stayed in Estonia for more than 183 days, which amounts to just 0.1% of visitors. It is very likely that the majority of these people from both countries work or have family relations in the neighbouring country.

Table 2.
Visits made between Estonia and Finland by duration, annual average for 2009-2011

Duration of visits	*Estonians to Finland*		*Finns to Estonia*	
	Number of visits	*%*	*Number of visits*	*%*
Transit	156 576	11.6	188 894	8.1
1 day	368 798	27.4	995 209	42.7
2 days	243 663	18.1	547 710	23.5
3 days	126 207	9.4	249 178	10.7
4 days	88 676	6.6	118 419	5.1
5 days	63 911	4.7	63 665	2.7
6 days	41 389	3.1	38 669	1.7
7 days	28 689	2.1	34 649	1.5
8 days	25 544	1.9	24 530	1.1
9 days	23 235	1.7	12 997	0.6
10+ days	180 487	13.4	57 564	2.5
Total	1 347 176	100	2 331 483	100.0

F. Frequency of trips

17. The annual frequency of trips per person (ID) also confirms that for the majority of visitors in both directions it is customary to make one or two trips per year (Table 3). Naturally, Finns dominate in making one-day trips: 72.2% of them go on one-day trips, whereas the share is only 49.9% for Estonians. Until up to 9 trips per year, the distribution of the frequency of trips is rather even between visitors of the two countries. In terms of determining labour migration, the focus is on people with 10 or more trips per year. Here, the share of persons who "often" travel from Estonia to Finland clearly stands out,

amounting to 10.6 % or 34,767 persons of the total number of travellers. The share of persons who travel "a lot" from Finland to Estonia is only 0.8% of the total number of travellers, i.e. 12,343 persons. Therefore, attention should be paid to people who work abroad particularly among persons who go to Finland 10 and more times per year. Where exactly the threshold should be set, depends on geographic distance and several other factors. Interviews and a phone use survey should probably be organised to help to fine-tune the methodology.

Table 3.

Visits made between Estonia and Finland by frequency of trips, annual average for 2009-2011

	From Estonia to Finland		*From Finland to Estonia*	
Frequency of visits	*Number of visitors*	*%*	*Number of visitors*	*%*
1	164 093	49.9	1 150 877	72,2
2	58 789	17.9	265 630	16,7
3	27 190	8.3	86 411	5,4
4	15 052	4.6	35 460	2,2
5	9 558	2.9	18 124	1,1
6	6 680	2.0	10 784	0,7
7	5 104	1.6	6 909	0,4
8	4 163	1.3	4 741	0,3
9	3 570	1.1	3 488	0,2
10+	34 767	10.6	12 343	0,8
Total	328 966	100	1 594 766	100,0

III. Temporary labour in Finland – Available data and limitations

A. Estonian sources and data limitations on temporary foreign labour in Estonia

18. Statistics Estonia can use two main sources to estimate the number and profile of those who are commuting for the purpose of work. First, population and housing census data offers sufficient data to estimate both the number of persons who were working abroad during a certain period (final weeks in 2011) as well as their personal characteristics. The disadvantage of this source is that it only gives information about 2011 and does not offer a continuous time series. The time series can be acquired from the Estonian labour force survey (LFS) but its main problem is the lack of sufficient respondents to describe more contrasted profiles. There would also be a problem of underestimating the number of those working abroad since it is difficult to reach them. Estimations based on the Estonian LFS can be used as a reference to the 2011 population and housing census; according to the survey the total number of people working abroad in the 1st quarter was 24,800. According to the 2011 population and housing census, the number of people working abroad was nearly 25,000. The number of people who worked in Finland was more than 15,000

according to the 2011 population and housing census, and the LFS showed similar results in the 4th quarter of 2011, in total 15,900 persons.

B. Data limitations for statistics on temporary foreign labour in Finland

19. Immigration to Finland has increased steadily and markedly since 1990's. As the volumes of immigration have increased, so has the need for information on a variety of issues related to immigration.

20. Statistical information on labour-based immigration is of particular interest for several administrative fields. Information on labour-based immigration is noticeably called for in formulating and monitoring policies and measures in economic and employment administration, in regional development, and in countering the black economy. However, statistics on labour-based immigration currently face severe limitations. Most importantly, official labour market and employment statistics are based only on population residing permanently in Finland.

21. For instance, Statistics Finland's annual register-based employment statistics, while yielding information on population's economic activity and employment, are based on permanent resident population. Those residing in Finland less than 12 months, or foreign nationals whose residence is attributed as temporary, regardless of the length of their stay, are not included in Statistics Finland's employment statistics.

22. Statistics Finland's LFS collects data on employment and unemployment. This data is based on a random sample, drawn twice a year from the Statistics Finland population database. The monthly sample consists of about 12,000 persons and data are collected with computer-assisted telephone interviews. Due to the nature of its sample construction, this survey does not reach those residing in Finland temporarily. In addition, since the data are collected via telephone interviews and due to possible language barriers, even permanently residing foreign nationals are at risk of being omitted from and under-represented in the LFS.

23. Statistics Finland's structural statistics on wages and salaries describe the numbers, hourly and monthly earnings, and the formation and distribution of employees' wages and salaries. The data derived from Statistics Finland's data on wages and salaries and is formed by combining data collected by employer organisations from their members with those from Statistics Finland's wage and salary inquiries for unorganised employers.

24. Structural statistics cover all public sector employees, but private enterprises employing less than five employees are not included in the data. In addition, agricultural (as well as forestry and fishery) enterprises are excluded from structural statistics. Due to these restrictions in data collection, Structural statistics leave out a cluster of small scale enterprises, many of which are run by self-employed immigrants, as well as enterprises in agriculture that typically rely on seasonal labour to a considerable extent.

25. The demand for information about temporary foreign labour is vast, and government agencies, in particular, have requested such data persistently. For instance, in 2004, the Finnish Parliament, while passing the directive on the right of citizens of the European Union to move and reside freely within the Member States, called for monitoring the effects of free movement on labour market. However, the revision and initiation of much legislation continuously requests reliable and accurate statistical data about temporary foreign labour.

26. The Programme for the Finnish Government 2011 establishes a definite objective to clarify and improve statistics and research activities concerning immigration. In its audit report *Työperäinen maahanmuutto* ("Labour- Based Immigration"), the National Auditing

Office notes that there are severe deficiencies in the availability and usability of statistics regarding temporary foreign labour, particularly from European Union member states.

27. The biggest obstacle for procuring statistics on temporary immigrant labour is the fact that foreign nationals do not acquire, for the most part, a Finnish social security number, which is imperative for register-based statistics production. The Finnish Aliens' Act is currently under revision and might eventually bring changes that will enable more fluent and comprehensive registering and, thus, produce statistics on temporary foreign labour. Proposed revisions to the Aliens' Act entail, for one, assigning a Finnish social security number to third country immigrants when they are issued a residence permit. Secondly, while the authority to assign social security number has earlier been restricted to magistrate offices, and more recently, to Office of Tax Administration, the revised Act would give the local police the right to assign social security numbers to EU-nationals after a three-month period of residence.

28. These revisions would expand the scope of foreign workers possessing a Finnish social security number and would therefore contribute to more consistent and reliable register-based statistics production. However, EU-residents staying in Finland less than three months would, even under the revised Act, remain unregistered and thus be without social security numbers. Further still, many EU-residents refrain from registration regardless, and many are not even aware that such an obligation exists in the first place.

29. Another profound deficiency is the fact that, currently, no government agency bears responsibility for statistics on labour-based immigration. In addition, while several agencies collect information about immigrants and immigration, no strategy to coalesce information from various registers and databases exists.

30. To respond to the need for and deficiencies in labour-based immigration statistics production, Statistics Finland has, from the initiation of the Ministry of Finance, set up a working group to investigate current limitations and deficiencies in registers, data, and statistics of labour-based immigration. The objective of the working group is to determine prerequisites for efficient and comprehensive statistics production of temporary immigrant labour, and consequently, present an action plan to meet these essentials.

31. Improving the registration processes of foreign workers, and developing processes to channel information about temporary foreign workers from enterprises to government agencies, are pressing needs. However, such measures aim to profoundly improve production of statistics on temporary foreign labour and can be lengthy processes. Therefore, prompt solutions for procuring data on labour-based immigration are in urgent need.

32. Increasing flows of commuting between Estonia and Finland is of particular interest for several policy-making authorities. However, free movement of labour within European Union territory sets definite limitations on data collecting of migrant labour, since border control mechanisms no longer yield data. Therefore, new and innovative approaches to data collecting must be developed.

IV. Discussion

A. Finnish perspective

33. Three main issues of concern were identified in relation to accessing mobile positioning data for the purpose of producing tourism statistics in Finland and these concerns will also potentially affect producing statistics about temporary labour migration.

34. First, the current Statistics Act obliges enterprises to provide data on their products and services but not on the clients consuming these products and services. Therefore, the current Statistics Act does not oblige MNOs to provide mobile positioning data to Statistics Finland.

35. Second, according to the Act on the Protection of Privacy in Electronic Communications identification data may only be processed by a natural person employed by or acting on behalf of a telecommunications operator.

36. Thirdly, the Office of Data Protection Ombudsman stated that the raw data, although anonymized, still constitutes personal data. The anonymization of data does not guarantee anonymity of the individual if the identifier individualising the subscriber connection is the same all the time, many observations are collected from the subscriber connection by means of location data and timestamps, and the time of data collection and the number of observations accumulated in this way are in no manner restricted. Thus, there arises a need for a mediator to process data into aggregated form on the premises of the operator. Such a mediator has to have access to the raw data maintained by the operator, as well as possess methodological competence and tools to process the raw data into aggregated data. This scenario involving a mediator is challenging technically, financially and in terms of organization.

37. As a conclusion there are two possible paths for Statistics Finland to gain access to the mobile positioning data:

(a) Voluntary basis: Operators process the data into aggregate form either themselves or by use of an intermediate party and deliver the aggregates to Statistics Finland.

(b) Legal basis: The Statistics Act should be updated to authorize Statistics Finland to obtain mobile positioning data from operators. In parallel, the Act on the Protection of Privacy in Electronic Communications should be updated to entitle the operators to provide data for Statistics Finland.

B. Can we trust the mobile data?

38. The data presented in this study is an alternative source for determining international mobility flows. Statistics are interesting and the data collection method is cost effective. Here we described only the first steps on developing such a segmented algorithm: a) duration of trips; b) frequencies of trips, in addition, we have done but not described here c) timing of trips; d) geographical origin and destination of trips. And the statistical offices discovered potential segments of travellers with mobile phones which are transnational commuters. They have different timing and duration of trips depending on occupation, season and economic situation in origin and destination.

39. Certainly these statistics also have problems: the total number of visitors was evaluated based on the market share of one large mobile operator. It must be emphasised that the study includes people that actively use their mobile phone while abroad, i.e., they have maintained an active connection with their homeland. Clearly there are many persons who use much cheaper local mobile phone service or do not use mobile phones at all while abroad. Nationality (country of origin) was determined by the country in which the phone is registered and usually people register their phones in a place with which they have a strong connection or where they stay longest. Certainly there is a need for a detailed phone use study in the future.

40. Using mobile positioning data in scientific research also has several shortcomings that we have to keep in mind when interpreting the results. One of the weaknesses of such quantitative statistical data is that we do not know the exact motivations behind the visits.

The most important question relates to sampling: Who are the persons who have phones? Are they using phones during travels? How often do visitors use phones in a foreign country? As roaming calls are expensive, it is likely that wealthy tourists and businessmen use their phones more often than less active people with a lower income (children, students, pensioners). This means that sampling issues are also related to lower income and age groups. Calling is also connected with cultural differences, such as calling regulations and traditions. Another problem that arises in the case of using mobile positioning data is its quantitative structure – there is information about the locations of calls (dots), but not on who is really making the calls, what kind of visit the person is on, and what kind of transportation he/she is using. The huge amount of quantitative data also poses a problem for data processing and cleaning; the databases are too large to enable using traditional software and data preparation options. From the positive side, getting mobile positioning data does not cause any extra burden and the production of the statistics can be done very fast. Given the limitation of traditional surveys and administrative register, a new alternative source that can give information about labour-based immigration is mobile positioning data. Mobile positioning data are a collection of information in a mobile operator's network on events that specify a specific subscriber's presence in time and space.

41. Estonia has been a leading country in the use of the mobile positioning data in various domains. This is because Estonia is one of the leading information technology (IT) countries with its e-government services and various public and private IT applications that create an environment that supports new technological development in many areas. Analyses based on mobile data in tourism, transportation and other public domains have been conducted by the University of Tartu, Positium and some other organisations already for several years.

42. It is clear that there are differences between the concepts used in the official statistics and what can be obtained from mobile positioning data. These discrepancies need to be taken into account when interpreting or comparing results. However, administrative data sources do not always measure what we want to measure, either. Assessment of accuracy, especially coverage issues, is very problematic for this data source because there are many components that contribute to the coverage bias and assessing all of them, separately or together, is a complex task and requires additional data sources. Some examples of coverage problems are: people not owning a mobile phone, people having many Subscriber Identity Module (SIM) cards, people changing their SIM to local SIM card when in other country, data available from selected MNOs only.

43. Passive mobile positioning data (the most likely data available from mobile network operators) contains data about events like making a call or sending a text message only. Due to the nature of passive mobile positioning data, the quality of the estimates based on this data source depend on changes in the telecommunication market, e.g. prices of the calls and text messages and the way individuals use their mobile phones. Mobile phone technology has developed very fast and people use mobile phones for much more than just calling and texting. It is quite likely that the added possibilities will change people's calling and texting habits and thus the content of data is changing too. Such change in the data will likely affect comparability over time. From the positive side, getting mobile positioning data does not cause any extra burden and the production of the statistics can be done very fast.

V. Conclusions

48. In this paper the main focus was to describe current results and methodology of passive mobile positioning and questions related this methodology. Advantages of this method over traditional methods used in border crossing statistics such as census, surveys

and counters, are better spatial and temporal resolution of data; a longer observation period, and the digital and automatic data collection. But also there are also problems in need of solutions. These problems concern legislative aspects, complete access to data, and need to develop methodology about measuring labour migration/temporary working. It is important to emphasize that pending questions would be related to both measuring mobility between countries and temporary working.

49. It is possible to bring out some further steps to move on with the process of developing. First, such segmentation needs to be properly finalised with a multi-layer analysis to consider all of these factors. Thus, from the aforementioned segmentation methods (frequency, duration, timing, geography), an algorithm could be compiled the help to identify the potential transnational commuters. Second, the result should be compared with national censuses and registers to determine the accuracy of such an algorithm of „passive" and anonymous positioning. It is vital to emphasise that migration always has two sides – a starting point and a destination – and migration methodology needs to consider the situation, knowledge and needs of both sides. Thirdly, we should organise a survey on phone use and positioning where people are interviewed and their phone use data is retrieved to be analysed in terms of the connections between their travel behaviour and IT use.

VI. References

Ahas, R. 2012. Virolaiset työntekijät ovat piilevä voimavara. Helsinkin Sanomat, 07.12.2012. http://www.hs.fi/paivanlehti/a1354767656107?jako=ab68c53e44198a132c2c0534436fbb42

Ahas, R., Silm, S., Järv, O., Saluveer E., Tiru, M. 2010. Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones . Journal of Urban Technology, 17(1): 3-27.

Ahas, R. Aasa, A., Roose, A., Mark, Ü., Silm, S. 2008. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. Tourism Management 29(3): 469–486.

Dahlman, E., Parkvall, S., & Skold, J. 2011. 4G: LTE/LTE-Advanced for Mobile Broadband, Elsevier.

EU, 2006. Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:EN:PDF

Järv, O., Ahas, R. and Witlox, F. 2014. Understanding monthly variability in human activity spaces: a twelve-month study using mobile phone call detail records. Transportation Research Part C: Emerging Technologies 38 (1): 122–135.

Massey, D.S., Goldring, L., Durand, H. 1994. Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities, American Journal of Sociology 99(6): 1492-1533.

Levitt, P.L., Jaworsky, B.N. 2007. Transnational Migration Studies: Past Developments and Future Treads, Annu. Rev. Sociol. 33:129–56.

Tyldum, G., Brunovskis A. 2005. Describing the unobserved: Methodological challenges in empirical studies on human trafficking, International Migration, 43(1/2):17-34.

Silm, S., Ahas, R. 2012. Spatial Mobility between Tallinn and Helsinki in Mobile Positioning Datasets. Statistical overview. Helsinki-Tallinn Transport and Planning Scenarios. Central Baltic INTERREG IV A Cross-Border Co-operation Programme, 66 p.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., Strogatz, S.H. (2010) Redrawing the map of Great Britain from a network of human interactions. PLoS ONE, 5(12): 1–6.

Smoreda, Z., Olteanu-Raimond A.M., & Couronné, T. 2013. In Zmud, J., et al. (Ed.) Transport Survey Methods: Best Practice for Decision Making. Emerald.

Zaiceva, A., Zimmermann K.F. 2008. Scale, diversity and determinants of labour migration in Europe. Oxford Review of Economic Policy, 24(3): 427-451.

Yuan, Y., Raubal, M., Liu, Y. 2012. Correlating Mobile Phone Usage and Travel Behavior -

A Case Study of Harbin, China. Computers, Environment and Urban Systems 36(2): 118-130.

Thanks to: Andres Oopkaup, Siim Krusell and Maiki Ilves of Statistics Estonia, Marjo Bruun of Statistics Finland, Rein Ahas of the University of Tartu, Estonia and Margus Tiru of Positium LBS, Estonia.

United Nations ECE/CES/2014/31

Economic and Social Council

Distr.: General
29 January 2014

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-second plenary session
Paris, 9-11 April 2014
Item 4 of the provisional agenda
Migration Statistics

Filling the 'migration gaps' — can research outcomes help us improve migration statistics?

Note by the Statistical Office of the European Union (Eurostat)

Summary

Migration statistics are known to be the most challenging part of demographic accounting. While continuing efforts to harmonise definitions and find new data sources have certainly improved the availability and quality of information on migratory flows, coherence/consistency in national and international migration data may still be the exception rather than the rule. Several research projects in recent years have helped by improving migration statistics, detecting alternative sources, sharing best practice or proposing estimation methods which address the quality and completeness of migration data. This paper explores the possible means offered by the research outcomes of dealing with the inherent uncertainty of migration data, and their potential use in the regular production and/or validation of official migration statistics.

The paper is presented for discussion to the Conference of European Statisticians seminar on migration statistics.

GE.14-20526

1420526

Please recycle

I. Introduction: the issue of gaps in mirror migration statistics

1. In the first twelve years of the new millennium, the population of Europe[1] saw growth of 2.5 %, of which three-quarters is attributable to migration. Developments in the European Union specifically were similar, with its population expanding by 3.6 % over the same period and net immigration of over 13 million persons. In several countries, migration was on a scale sufficient to cancel out the natural decline in the population. Correct estimates of migration flows and stocks are thus of paramount importance.

2. Population size and structure can no longer be properly gauged without an accurate estimate of migration. The role that migration has assumed in European population dynamics imposes a 'statistical duty' of coherence between the factors in the fundamental demographic balance equation: $P(t+1) = P(t) + B(t) - D(t) + I(t) - E(t)$, i.e. the population at the end of the year is the sum of the population at the beginning of the year, plus natural change expressed by the difference between live births (B) and deaths (D), plus the net number of migrants entering (I, immigrants) or leaving (E, emigrants). The population cannot be computed if migration flows are not measured properly.

3. Even if coherence is achieved within the demographic balance of a country, comparability at international level may still be a challenge. Unlike vital events or population data, logically, migration statistics recorded in the countries from/to which the flows originated/went should correspond. Common sense dictates that national sets of migration figures should be at least approximately consistent. In simple terms, the number of immigrants recorded in country A as coming from country B should match the number of emigrants from country B heading for country A. Unfortunately, this is hardly ever the case.

4. The issue of mismatching mirror statistics in migration is a longstanding one. As long ago as 1975, the Conference of European Statisticians (CES) noticed the discrepancies between immigration and emigration flows involving the same pair of countries in an origin/destination matrix of migratory flows in 1972. A number of reasons for these migration 'gaps' were identified at the time: differing definitions of 'immigrants' and 'emigrants', different statistical coverage, differences in the statistical treatment of temporary migrants, a lack of consistency in classifying migrants by country of origin/destination, and differences between countries in the data sources and methods of collection used (Kelly 1987).

5. These observations triggered a range of activity, in both the research and statistical fields. Statisticians have sought to improve the comparability of migration statistics, adapting national to internationally recommended definitions (to varying degrees), exchanging mirror data with correspondent countries and/or carrying out bilateral studies. Researchers have developed a variety of methods to overcome inconsistencies in migration data. However, over 40 years later, migration statistics still present significant inconsistencies. The latest statistical evidence from selected European countries shows that the number of immigrants from country A measured in destination country B can differ from the number of emigrants to country B measured in country of origin A. This also

[1] For statistical convenience, 'Europe' refers here to the 44 countries for which data were available and does not reflect any opinion on the part of the European Commission concerning the legal status of any country, territory or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

occurs between countries which should in principle provide harmonised statistics of migration flows.

6. Although much of the focus is now on migrants' socio-economic characteristics, measuring their integration and links between migration and the labour market, this needs to be based on an accurate idea of immigrant/emigrant numbers – otherwise, any secondary information will be distorted. This note explores the opportunities and challenges involved in applying the results of research on migration statistics in regular data production. Section II briefly shows (without entering into policy issues) the specific analytical relevance of migration statistics for the European Union (EU), which has made the EU a popular case study. Section III summarises a number of outcomes from academic and statistical research and major projects supported by the EU in recent years. Section IV opens a discussion on some potential issues regarding the adoption of estimation methods in migration statistics. A few concluding remarks are made in the last section.

II. The European Union as a case study for migration statistics

7. Migration statistics have played a key role in numerous policy actions in recent years. Acknowledgement of immigration as an important resource in a country's population dynamics and a greater focus on the implications of emigration have helped raise awareness of the statistical information available. Among other action in this field,[2] in 2008 the Council of the European Union formally established the European Migration Network (EMN), which is geared to supporting policy-making by providing Community institutions and Member State authorities with up-to-date, objective, reliable and comparable information on migration and asylum.[3] The latest EMN conference called for a stronger evidence base, including more analyses of migration statistics (EMN 2013).

8. The EU regulations on migration statistics[4] are a milestone in this field. Since their entry into force, a large number of European countries have worked to deliver more harmonised migration statistics. The regulations were the outcome of a two-decade-long process launched in response to rapidly growing interest in migration statistics and realisation of the need for harmonisation at EU level (Poulain 1991).

9. However, as noted above, such efforts are not sufficient to provide completely consistent and high-quality information on migration flows in the EU. Acknowledging the inherent difficulties in measuring such phenomena, the regulations provided for the application of 'scientifically based and well documented statistical estimation methods' (Article 9 of Regulation (EC) No 862/2007). The combination of policy relevance and endorsement in EU law may well have been a factor in (if not the root cause of) the resurgence of scientific work in the area of migration, with many recent contributions focusing on the problems of statistical estimation in the EU context.[5]

10. The peculiarity of the EU means that some observers regard intra-EU migration as 'internal' rather than international. This provides additional justification for the adoption of

[2] For more information on EU immigration policies, see http://ec.europa.eu/dgs/home-affairs/what-we-do/policies/immigration/index_en.htm and http://ec.europa.eu/dgs/home-affairs/what-we-do/policies/international-affairs/global-approach-to-migration/index_en.htm.

[3] Council Decision 2008/381/EC establishing a European Migration Network (OJ L 131, 21.5.2008, p. 7).

[4] Regulation (EC) No 862/2007 on Community statistics on migration and international protection (OJ L 199, 31.7.2007, p. 23) and its implementing measures.

[5] This is not to say that similar problems do not apply in other parts of the world (see, for example, Van Hook and Bachmeier 2013, Judson and Swanson 2011, Massey and Singer 1995).

estimation methods hitherto applied in inter-regional migration models. Also, the EU may intuitively be considered a 'migration system',[6] or set of migration systems, which also makes it interesting from a theoretical point of view.

11. The particular impact of the economic crisis on some European countries provides a further opportunity for EU-specific analyses using migration data. As shown by Herm and Poulain (2012), it is essential to ensure that time series are consistent – a real challenge considering the changes in national methodology, definitions, etc. that have characterised this field over the years. Proper assessment of migration patterns and trends requires the reconstruction of reliable time series of past immigration and emigration.

III. A short review of research on European migration statistics

12. Several projects on migration statistics have been funded by the European Commission, particularly under the 5th, 6th and 7th Research Framework Programmes; these include (in chronological order):

(a) January 2001-October 2002: *the Comparing National Data Sources in the Field of Migration and Integration* (COMPSTAT) project[7] took initial steps to overcome the lack of comparable data in 10 European countries (Austria, Belgium, the Czech Republic, Germany, Hungary, Italy, the Netherlands, Norway, Poland and Switzerland);

(b) April 2004-August 2005: the *Towards Harmonised European Statistics on International Migration* (THESIM) project[8] was aimed at supporting implementation of the EU regulation on migration statistics and helping countries comply with their obligations. The outcomes were also published in Poulain *et al.* (2006);

(c) January 2007-December 2009: the Eurostat-funded *Migration Modelling for Statistical Analysis* (MIMOSA) project[9] aimed to develop methods for reconciling differences in European countries' international migration statistics. The project produced estimates of migration flows and population stocks for the 31 EU and EFTA countries for 2002-07;

(d) March 2007-August 2009: the *Promoting comparative quantitative research in the field of migration and integration in Europe* (PROMINSTAT) project[10] produced an online database containing essential technical information on migration and integration for 29 European countries (EU-27 plus Norway and Switzerland);

(e) September 2007-August 2009: the *Irregular Migration: Counting the Uncountable. Data and Trends Across Europe* (CLANDESTINO) project[11] aimed to provide an inventory of data and a new method for evaluating estimates of irregular migration (stocks and flows) in selected EU countries (Austria, Germany, Greece, Hungary, Italy, the Netherlands, Poland, Portugal, Slovakia, Spain and the United Kingdom). It set up

[6] Although not as yet the subject of commonly accepted definition and application, the term 'migration system' basically refers to a network of places linked by flows and counter-flows of people (see, for example, Bakewell 2013).

[7] The project website is no longer available; for some information, see http://www.efms.uni-bamberg.de/prsteuce.htm.

[8] http://www.uclouvain.be/en-7823.html.

[9] http://mimosa.gedap.be/.

[10] http://www.prominstat.eu.

[11] http://clandestino.eliamep.gr/.

a database on irregular migration,[12] hosted by the Hamburg Institute of International Economics (HWWI);

(f) Funded by the NORFACE Research Programme on Migration, the *Integrated Modelling of European Migration* (IMEM) project[13] aims to overcome the limitations of the various data sources, including inconsistent availability, definitions and quality, by applying Bayesian statistical methods.

13. Many other EU-funded projects have addressed (albeit not always as their primary objective) the availability and quality of migration data, and a number of databases have either been set up specifically or happen to give insights on determinants of migration which could also be used for estimation purposes (e.g. Kim and Cohen 2010). The following is a non-exhaustive list (in chronological order):

(a) October 2005-September 2007: the *Patterns of migration in the new European Borderlands: an assessment of post-enlargement migration trends in NIS border countries (Armenia, Belarus, Georgia, Moldova and Ukraine)* project, funded by the INTAS Programme and based on survey information in the five countries in question, explored the factors which influence migration duration and the choice of destination region (for an outcome of this research, see Danzer and Dietz 2008);

(b) January 2007-June 2009: the *Mediterranean and Eastern European Countries as New Immigration Destinations in the European Union* (IDEA) project[14] had as one of its main objectives the preparation of a model of future migratory trends in selected parts of the EU, based on a collection of data on population stocks and flows in the countries in question in 1985-2004. A book edited by Okólski (2012) sets out some of the outcomes;

(c) October 2008-January 2013: the *Migration between Africa and Europe* (MAFE) project[15] focused on migration flows between Europe (Belgium, France, Italy, the Netherlands, Spain and the UK) and sub-Saharan Africa (Senegal, the Democratic Republic of Congo and Ghana) and collected unique data on migrant characteristics and behaviour;

(d) January 2010-December 2014 (ongoing): the *Determinants of International Migration* (DEMIG) project[16] announces, in its call for papers for a conference at Oxford University in September 2014, empirical analyses of large-scale country-to-country migration flow databases, total in- and outflows, migration policies and visa requirements;

(f) June 2012-November 2014 (ongoing): the *Managing Migration and its Effects in South-East Europe — Transnational Actions Towards Evidence Based Strategies* (SEEMIG) project[17] has among its outputs an analysis of existing migratory data production systems and data sources, and a database of immigrant flows and stocks;

(g) Funded by the NORFACE Research Programme on Migration, the *Theorising the Evolution of European Migration Systems* (THEMIS) project[18] follows migration trajectories from certain regions in three origin countries (Brazil, Morocco and Ukraine) to selected cities in four destination countries (the UK, Norway, the Netherlands and Portugal).

[12] http://irregular-migration.net//index.php?id=157.

[13] http://www.norface-migration.org/currentprojectdetail.php?proj=3 or http://www.norface.net/526.

[14] http://www.idea6fp.uw.edu.pl/.

[15] http://www.mafeproject.eu/.

[16] http://www.imi.ox.ac.uk/projects/demig or www.migrationdeterminants.eu.

[17] http://www.seemig.eu/.

[18] http://www.imi.ox.ac.uk/projects/themis. See also http://www.norface.net/526 and http://www.norface-migration.org/currentprojectdetail.php?proj=11.

13. A number of contributions have been published in international scientific journals, in some cases independently of the above projects. The following is a (non-exhaustive) list of contributions over the past twenty years, most focusing on European countries and, in particular, migratory flows involving EU Member States:

(a) Poulain (1993) develops a basic idea proposed earlier (Poulain and Wattelar 1983) which involves applying correction factors to overcome migration gap problems in mirror statistics. Variants of this approach are developed by Abel (2010) and De Beer *et al.* (2010), the latter subsequently followed up by Raymer *et al.* (2011) with the estimation of missing data and a breakdown by age and gender.

(b) Willekens (1994) provides a first review of available methods in demography, geography and statistics which could have been applied to set up a consistent and coherent database on international migration, including a call for the inclusion of judgmental data. The same scholar presents various probability models from a statistical perspective, with particular attention to life history data (Willekens 1999a) and a probabilistic method for estimating migration from incomplete data using the 'expectation-maximisation' algorithm (Willekens 1999b).

(c) Nowok and Willekens (2011) propose a probabilistic model for migration processes, focusing on an assessment of the impact that different migration/migrant definitions have on measurements.

(d) DeWaard *et al.* (2012) extend an approach developed under the MIMOSA project to estimate international migration flows between 28 European countries (EU-27 and Norway) in 2003-07, additionally identifying three migration systems in the region.

(e) Abel (2013) proposes a method for estimating migration flows using place-of-birth data, linking migration flows to stocks.

(f) In the framework of another EU-funded project,[19] Cuaresma *et al.* (2013) develop a new method for estimating bilateral migration flows using an econometric specification of a gravity model based on net migration.

(g) Wiśniowski *et al.* (2013) show how experts' opinions can be incorporated in a Bayesian model to improve migration flow statistics; this can also be used to cover irregular migration.

(h) Raymer *et al.* (2013) propose a method for estimating migration flows that incorporates covariate information, experts' opinions and accuracy of data collection systems, and is accompanied by measures of uncertainty.

(i) Last, but definitely not least, are books written by Rogers *et al.* (2009), on migration estimation methods, and edited by Raymer and Willekens (2008), the latter specifically addressing the data harmonisation challenges presented by the EU Regulation.

14. Other contributions from national statistical offices have mainly addressed ways to improve the primary information, for instance by looking at the accuracy of national data sources (e.g. ISTAT-Italy 2008), reviewing definition and measurement issues (e.g. Statistics Austria 2008), comparing migration data systems in detail (e.g. Statistics Netherlands and Statistics Sweden 2010) or exchanging (ultimately micro-)data (INE-Spain 2013).

15. At international level, the practice of data exchange has been supported by the United Nations Economic Commission for Europe (UNECE), which has also issued guidelines on this (UNECE 2010), and Eurostat, which has set up a secured web repository for countries to exchange migration data before official release and document best

[19] *Welfare, Wealth and Work for Europe* (WWW for Europe): http://www.foreurope.eu/.

practices. Discussion has taken place on the impact of different definitions of migration (UNECE Task Force 2012) and more recently on new forms of migration, e.g. in the recently established UNECE Task Force on Circular Migration.[20] Finally, alternative data sources have been explored by the Suitland Working Group.[21]

IV. Discussion

16. The previous section highlighted just part of the wealth of methods now available to official statisticians for improving migration statistics. Although scientists must continue the quest for ever better methods, statistical offices responsible for producing and/or disseminating migration data can no longer ignore those already available. Nevertheless, the potential seems to date to have been under-exploited. One possible reason is insufficient awareness (as yet) of the progress achieved, or a sense that this is insufficiently conclusive. Also, implementation of some new methods in regular production may require particular technical skills, which may not be widespread. Thirdly, several of the methods, particularly those referring to the origin-destination matrix of migration flows, are better tackled at international level.

17. Other factors may hamper the implementation of research outcomes in regular migration data production. For the purposes of the current discussion, we can apply a (somewhat fuzzy) distinction between *statistics* and *estimates*, the former broadly representing the product of an (as far as possible) exhaustive compilation of records from primary data source(s), the latter the outcome of probabilistic/statistical models, possibly combining information from various sources. Migration data from sampling surveys would fall into this second category.

18. The United Nations Statistics Division (UNSD) and Population Division (UNPD) operate on the basis of this distinction. The UNSD collects official statistics from all countries, which the UNPD then uses to produce internationally consistent estimates of population and related components of change (including migration). However, the timeliness and granularity of the information made available by the UNPD may not be such as to meet the needs of regional policy-makers (e.g. in the EU).

19. Even when an estimation model is adopted, there are a number of potentially far-reaching consequences, summed up by the following questions:

(a) Should countries adopt the migration estimates as national figures?

(b) Should countries adjust their population figures accordingly?

(c) Can migration/population *estimates* be used for 'official' purposes, instead of 'original' migration/population *statistics*?

20. Whatever the national position on these questions (which may depend on domestic priorities), data comparability and completeness are a must for any international analysis. Estimates, in particular those relating to the origin-destination matrix of migration flows, are also an invaluable tool for data validation and quality assessment. However, it must be stressed that the adoption of any estimation method does not justify relaxing the commitment constantly to improve data sources and derived statistics.

21. In practice, where migration gaps are due to differences of definition, an estimation method that does not take this into account is 'forcing' consistency between estimates and

[20] For terms of reference, see http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2013/15Add.1_-ToR_TF_migration.pdf.

[21] For more information, see http://www.unece.org/stats/groups/suitland/suitland.html.

mixing dissimilar measures. The outcome can be seen as referring to an undefined general concept of migration, of which the definitions are thus considered an approximation. This highlights once more the importance of continuous effort to produce data based on definitions that are as comparable as possible.

22. Some estimation methods may require further refinement/development before being directly applicable. In addition, various breakdowns, e.g. by age, sex, country of birth and citizenship, are of key interest, so consistency is important not only for top-level aggregated figures, but for the whole set of data.

23. A risk in neglecting the possibilities offered by estimation methods is that, in the absence of adequate migration data, users may turn to third-party estimates that do not necessarily meet statistical/scientific requirements; this may trigger potentially embarrassing disputes. If estimates have to be used, a statistical office is well advised to master them, also for the sake of complying with its overall duty to produce accurate information for the society it serves.

24. Ideally, the application of estimation methods should be managed at international level. Even if a country adopts a particular method, the outcomes may not match those from another method applied in another country – there would be a mismatch of estimates rather than statistics! It seems, therefore, that action is needed at international level to coordinate the implementation of estimation methods in migration statistics.

V. Concluding remarks

25. The policy relevance and intrinsic international nature of migration call for national statistical offices to continue working towards the full harmonisation of relevant data. In view of the role now played by migration in national population dynamics, such efforts should be seen in the framework of harmonising population statistics in general: migration statistics can no longer be regarded as separate or negligible.

26. While countries are undeniably entitled to adopt particular approaches for their own purposes, in a globalised world there is a clear need to move towards a *single* concept of 'population' for international statistical purposes, including a common definition of 'migration'. The perpetual debate as to how to define 'population' ('usually resident' vs. 'registered' vs. 'legal population', etc.), the challenges engendered by greater mobility and inter-connectedness, and the varying concepts of 'population' underlying national statistical (and sometimes legal) systems may mean that a paradigm shift is required (see, for example, Eurostat 2013).

27. After forty years of inconsistent migration data at international level, a move forward is now called for. The research community has been suggesting solutions for some time and, although estimation methods can hardly replace accurate data sources, statisticians have a valuable set of tools at their disposal, for now mainly as regards numbers of immigrants and emigrants by country of origin/destination. Although some techniques may be considered insufficiently mature as yet, mirror tables are an easy-to-build indicator of the international consistency of migration statistics, and consistent sets produced using estimation methods can be a precious source of support for data validation and assessment. In migration statistics, *filling the gaps* in the mirror tables involves *bridging the gaps* between statisticians and researchers.

28. Given its official role in collecting migration data from EU Member States and the benefits of an international approach, Eurostat might now consider investigating the technical pros and cons, and performance, of the various methods of migration estimation, in view of their regular application.

VI. References

Abel, G.J. (2010): "Estimation of international migration flow tables in Europe". *Journal of the Royal Statistical Society*, 173(4):797-825.

Abel, G.J. (2013): "Estimating global migration flow tables using place of birth data". *Demographic Research*, 28(18):505-546.

Bakewell, O. (2013): "Relaunching migration systems". *Migration Studies*, Advance Access. First published online November 15, 2013 doi:10.1093/migration/mnt023.

Cuaresma, J.C., M. Moser and A. Raggl (2013): "On the Determinants of Global Bilateral Migration Flows". Working Paper no.5 of the project "WWWforEurope - Welfare, Wealth and Work for Europe".

Danzer, A.M., and B. Dietz (2008): "Economic Migration, Networks and Human Capital Transferability from the New European Borderlands. A Comparison of Five Eastern European Countries". Proceedings of the German Development Economics Conference, Zurich, 30-31 May 2008.

De Beer, J., J. Raymer, R. van der Erf and L. van Wissen (2010): "Overcoming the Problems of Inconsistent International Migration Data: A New Method Applied to Flows in Europe". *European Journal of Population*, 26:459-481.

DeWaard, J., K. Kim and J. Raymer (2012): "Migration Systems in Europe: Evidence From Harmonized Flow Data". *Demography*, 49:1307-1333.

EMN - European Migration Network (2013): "The Agenda for Growth in the EU : Why Migration Matters ? – Summary and Final Conclusions". European Migration Network Conference, Vilnius (LT), 5-6 November 2013.

Eurostat (2013): "On a New Definition of Population for Statistical Purposes". Paper for the 15th Meeting of the UNECE Group of Experts on Population and Housing Censuses, Geneva, 30 September - 3 October 2013.

Herm, A., and M. Poulain (2012): "Economic Crisis and International Migration. What the EU Data Reveal?" *Revue Européenne des Migrations Internationales*, 28(4):145-169.

INE-Spain (2013): "Exchange of information between EU Member States to improve population statistics. A proposal from INE-Spain". Paper for the meeting of the Eurostat Working Group on Population Statistics, Luxembourg, 16 - 17 December 2013.

ISTAT-Italy (2008): "Statistics on migration in Italy: a reassessment of sources and methods". Paper for the CES Seminar on "Measuring population movement and integration in a globalized world", Geneva, 10 - 12 June 2008.

Judson, D.H, and D.A. Swanson (2011): "Estimating Characteristics of the Foreign Born by Legal Status - An Evaluation of Data and Methods". *SpringerBriefs in Population Studies*, Volume 2.

Kelly, J. (1987): "Improving the Comparability of International Migration Statistics: Contributions by the Conference of European Statisticians from 1971 to Date". *International Migration Review*, Special Issue "Measuring International Migration: Theory and Practice", 21(4):1017-1037.

Kim, K., and J.E. Cohen (2010): "Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity". *International Migration Review*, 44(4):899-932.

Massey, D.S., and A. Singer (1995): "New Estimates of Undocumented Mexican Migration and the Probability of Apprehension". *Demography*, 32(2):203-2013.

Nowok, B., and F. Willekens (2011): "A Probabilistic Framework for Harmonisation of Migration Statistics". *Population, Space and Place*, 17:521-533.

Okolski, M., ed. (2012): "*European Immigrations: Trends, Structures and Policy Implications*". Amsterdam University Press.

Poulain, M. (1991): "Un projet d'harmonisation des statistiques de migration internationale au sein de la Communauté Européenne". *Revue Européenne des Migrations Internationales*, 7(2):115-138.

Poulain, M. (1993): "Confrontation des Statistiques de Migrations Intra-Européenne : Vers plus d'Harmonisation ?" *European Journal of Population*, 9:353-381.

Poulain, M., N. Perrin and A. Singleton, eds. (2006): "*THESIM – Towards Harmonised European Statistics on International Migration*". Presses Universitaires de Louvain.

Poulain, M., and C. Wattelar (1983): "Les migrations intra-européennes : à la recherche d'un fil d'Ariane au travers des 21 pays du Conseil de l'Europe". *Espaces, populations, sociétés*, 2:11-26.

Raymer, J., J. de Beer and R. van der Erf (2011): "Putting the Pieces of the Puzzle Together: Age and Sex-Specific Estimates of Migration amongst Countries in the EU/EFTA, 2002-2007". *European Journal of Population*, 27:185-215.

Raymer, J., and F. Willekens, eds. (2008): "*International Migration in Europe –Data, Models and Estimates*". Wiley.

Raymer, J., A. Wiśniowski, J.J. Forster, P.W.F. Smith and J. Bijak (2013): "Integrated Modeling of European Migration". *Journal of American Statistical Association*, 108(503):801-819.

Rogers, A., J. Little and J. Raymer (2009): "*The Indirect Estimation of Migration – Methods for Dealing with Irregular, Inadequate, and Missing Data*". Springer.

Statistics Austria (2008): "Migration statistics: conceptualisation and measurement in the European Union". Paper for the CES Seminar on "Measuring population movement and integration in a globalized world", Geneva, 10 - 12 June 2008.

Statistics Netherlands and Statistics Sweden (2010): "Comparing international data on migration between Sweden and the Netherlands". Paper for the Joint UNECE/Eurostat Work Session on Migration Statistics, Geneva, 14-16 April 2010.

UNECE (2010): "Guidelines for Exchanging Data to Improve Emigration Statistics". Prepared by the Task Force on "Measuring Emigration Using Data Collected by the Receiving Country". United Nations, Geneva.

UNECE Task Force (2012): "Analysis of international migration estimates using different length of stay definitions". Paper for the Joint UNECE/Eurostat Work Session on Migration Statistics, Geneva, 17-19 October 2012.

Van Hook, J., and J.D. Bachmeier (2013): "How well does the American Community Survey count naturalized citizens?" *Demographic Research*, 29(1):1-32.

Willekens, F. (1994): "Monitoring International Migration Flows in Europe - Towards a Statistical Data Base Combining Data From Different Sources". *European Journal of Population*, 10:1-42.

Willekens, F. (1999a): "Probability models of migration: complete and incomplete data". *SA Journal of Demography*, 7(1):31-42.

Willekens, F. (1999b): "Modeling Approaches to the Indirect Estimation of Migration Flows: From Entropy to EM". *Mathematical Population Studies*, 7(3):239-278.

Wiśniowski, A., J. Bijak, S. Christiansen, J.J. Forster, N. Keilman, J. Raymer and P.W.F. Smith (2013): "Utilising Expert Opinion to Improve the Measurement of International Migration in Europe". *Journal of Official Statistics*, 29(4):583-607.